Perturbation of hyperbolic polynomials and related lifting problems

Armin Rainer

Fakultät für Mathematik, Universität Wien, Nordbergstrasse 15, A-1090 Wien, Austria

E-mail address: `armin.rainer@univie.ac.at`

2000 *Mathematics Subject Classification.* 26C10, 22E45, 20F55, 57S15, 57S17

Key words and phrases. hyperbolic polynomials, perturbation, differentiable parameterizations of roots, proper G-manifolds, compact Lie group representations, invariants, differentiable liftings

The author was supported by the Austrian Science Fund (FWF), Grants P14195, P17108, J2771, and P22218.

Preface

The notion of hyperbolic polynomials originates from the theory of partial differential equations. It probably appears the first time in the fundamental paper [**Går51**] published in 1951 due to L. Gårding. This notion of hyperbolicity reflects an algebraic condition necessary for the well-posedness of a Cauchy problem: Let $H \subseteq \mathbb{R}^n$ be a half space defined by an inequality $\langle x \mid N \rangle \geq 0$. Let $P(\xi)$ be a polynomial with complex coefficients in n variables $\xi_1, \dots, \xi_n$, and let $P(D)$ be the differential operator given by replacing ξ_j by $D_{x_j} = -i\frac{\partial}{\partial x_j}$. The Cauchy problem for $P(D)$ in H with homogeneous Cauchy data is to find a solution of the equation

$$P(D)u = f$$

in $\mathbb{R}^n$ with $\operatorname{supp} u \subseteq H$ when f is given with $\operatorname{supp} f \subseteq H$. One can show that $f = 0$ then implies $u = 0$ if and only if $P_m(N) \neq 0$, where P_m is the principal part of P. Moreover, it turns out that if the equation $P(D)u = f$ has a solution $u \in \mathscr{D}'(\mathbb{R}^n)$ with support in H for every $f \in C_0^\infty(H)$ and $P_m(N) \neq 0$, then there exists a number τ_0 such that

$$P(\xi + i\tau N) \neq 0$$

for $\xi \in \mathbb{R}^n$ and $\tau < \tau_0$. If these hypothesis are satisfied then the equation $P(D)u = f$ has a unique solution $u \in C^\infty(\mathbb{R}^n)$ with $\operatorname{supp} u \subseteq H$ for every $f \in C^\infty(\mathbb{R}^n)$ with $\operatorname{supp} f \subseteq H$. See [**Hör83b**, chapter XII]. A polynomial P is called hyperbolic with respect to the real vector N, if $P_m(N) \neq 0$ and if there exists a $\tau_0 > 0$ such that $P(\xi + i\tau N) \neq 0$ for $\xi \in \mathbb{R}^n$ and $\tau < \tau_0$.

It is not hard to see that a homogeneous polynomial P is hyperbolic with respect to N if and only if $P(N) \neq 0$ and the equation

$$P(\xi + \tau N) = 0$$

has only real roots when ξ is real. Homogeneous hyperbolic polynomials enjoy some interest, apart from partial differential equations, in the context of convex analysis, e.g. [**BGLS01**].

The characterization of homogeneous hyperbolic polynomials leads us to the related notion of hyperbolic polynomials we are going to deal with primarily in this book: We shall call a monic polynomial in one variable

$$P(x) = x^n - a_1 x^{n-1} + \cdots + (-1)^n a_n$$

with real coefficients a_i hyperbolic if all its roots are real.

The study of polynomials in one variable, of their zeros and critical points, is a classical topic in mathematics. There is the algebraic viewpoint on the one hand, when considering the polynomials as algebraic expressions in one unknown and their zeros as roots of an equation, and the analytic approach on the other hand, viewing them as functions of a particular kind with excellent analytic properties. Classical results of algebraic nature are for instance the fundamental theorem of algebra, the Gauss-Lucas theorem, the Budan-Fourier theorem, Sturm's theorem, Cauchy's index theorem, etc., all of them dealing with the number and location of the zeros, while analytical interest in polynomials concentrates on their extremal properties, such as in the work of Chebyshev, in Weierstrass' approximation theorem, or Bernstein's inequalities. It should be mentioned that most results allocated to the algebraic branch do not dispense with some analytic nature. An excellent treatment of both branches, algebraic and analytic, and their interrelation, as well as rich bibliographical references, is provided by the monograph [**RS02**].

Hyperbolic polynomials in one variable, in particular, the properties and the structure of spaces of hyperbolic polynomials of fixed degree play a role in singularity theory, for instance. We mention in this context the work of V.I. Arnold and his school.

In this book we are principally interested in the question how the roots of a polynomial behave if the polynomial is perturbed. More precisely: Let us assume that the coefficients a_i of P depend on a parameter t in some regular way:

$$P(t)(x) = x^n - a_1(t)x^{n-1} + \cdots + (-1)^n a_n(t).$$

We search for regular functions $x_1(t), \ldots, x_n(t)$ which parameterize the roots of $P(t)$ for any t. We shall be mostly interested in the restriction of that problem to hyperbolic polynomials $P(t)$ for two reasons. Firstly, this perturbation problem for hyperbolic polynomials has the most important applications, and, secondly, without restriction one can only expect very little as illustrated by the following example: The roots of $x^2 - t = 0$, $t \in \mathbb{R}$, cannot be Lipschitz continuous near $t = 0$, despite the fact that the dependence of the polynomial on the parameter t is real analytic.

Moreover, we will restrict ourselves basically to the one dimensional real case, i.e., we will suppose that the parameter t runs through some interval in $\mathbb{R}$. This is because for higher dimensional parameters there are 'only' negative results: Not even continuous parameterizations of the roots are in general possible. Continuity of the set of roots as a whole as well as partial differentiability is true, but not total differentiability: The roots of the equation $x^2 - (t_1^2 + t_2^2) = 0$ are not totally differentiable at $(t_1, t_2) = (0, 0)$. However, in the hyperbolic case, given that the coefficients of P depend on $t \in \mathbb{R}^m$ in a C^n way, ordering the roots increasingly provides a continuous parameterization which turns out to be locally Lipschitz. More can in general not be expected.

The problem of parameterizing regularly the roots of a regular curve of hyperbolic polynomials of fixed degree naturally is rooted in the perturbation theory for linear operators. The main question in this theory is how the eigenvalues and eigenvectors change with the operator. If the operators act in a finite dimensional space, its eigenvalues are exactly the roots of the associated characteristic polynomial. Moreover, if the operators are selfadjoint, then all eigenvalues are real, and, thus, the characteristic polynomial is hyperbolic. The perturbation theory for selfadjoint finite dimensional operators has been developed by F. Rellich in [**Rel37**], [**Rel42**], [**Rel69**]; see also [**Kat76**]. The study of perturbations of hyperbolic polynomials may then be considered as a generalization of that question, namely, it amounts to omitting the restriction to hyperbolic polynomials which arise as characteristic polynomials of selfadjoint operators. Rellich himself proved in [**Rel37**] that if the coefficients $a_i(t)$ of the polynomial $P(t)$ are real analytic functions in one variable t, then one can choose the roots of $P(t)$ in a real analytic way. It turns out that in general the omission of that restriction involves a weakening of the attainable results.

A second motivation for studying perturbations of hyperbolic polynomials originates again from the theory of partial differential equations. The fundamental paper [**Bro79**] due to M.D. Bronshtein, in which basically was proved that a C^n curve of monic hyperbolic polynomials of degree n admits a differentiable parameterization of its roots with locally bounded derivative, enabled the same author to show in

[**Bro80**] the well-posedness of the hyperbolic Cauchy problem

$$P(x, D)u(x) = f(x)$$

with non-constant coefficients in Gevrey space. A different approach which led to a slightly more general result than in [**Bro79**] was given by S. Wakabayashi in [**Wak86**]. This second approach is more conceptual and shorter.

Bronshtein's result on the existence of differentiable parameterizations of the roots with locally bounded derivative seems to be not well known outside the community of specialists in hyperbolic partial differential equations. Its proof in [**Bro79**] is rather sketchy and was not really accepted (or, more likely, not understood) by the specialists. What makes is quite delicate is that it is not possible to give any explicit bound for the derivative of the roots, for instance in terms of bounds for derivatives of the coefficients. One concern of this book is to give a detailed, understandable, and complete elaboration of Bronshtein's proof. Furthermore, we provide also an accurate and exhaustive presentation of Wakabayashi's approach.

The investigation of perturbations of hyperbolic polynomials initiated by Rellich has henceforth resulted in a rather complete picture of the situation. Let us summarize the most relevant results (in their essential form, sometimes not as general as originally stated): Consider a curve

$$P(t)(x) = x^n - a_1(t)x^{n-1} + \cdots + (-1)^n a_n(t) \quad (t \in \mathbb{R})$$

of hyperbolic polynomials.

- If the coefficients a_i are real analytic, there exist real analytic parameterizations of the roots of P. (Rellich [**Rel37**], 1937).
- The square root of a non-negative C^∞ function in m variables is in general at most a C^1 function. (Glaeser [**Gla63b**], 1963; Dieudonné [**Die70**], 1970).
- If all a_i are of class C^n, then there exists a differentiable parameterization of the roots of P with locally bounded derivative. (Bronshtein [**Bro79**], 1979).
- If all a_i are of class C^{2n}, any differentiable choice for the roots is actually C^1. (Mandai [**Man85**], 1985).
- If all a_i are of class $C^{k,\alpha}$, where $0 < \alpha \le 1$, then on any open bounded interval I the increasingly ordered roots of P satisfy a Hölder condition with exponent $\min\{1, \frac{k+\alpha}{n}\}$. (Wakabayashi [**Wak86**], 1986).
- If all a_i are of class C^∞ and no two of the increasingly ordered continuous roots meet of infinite order of flatness, then there exist C^∞ parameterizations of the roots and any two choices differ by a permutation. (Alekseevsky, Kriegl, Losik, Michor [**AKLM98**], 1998).
- If all a_i are of class C^{3n}, there exists a twice differentiable parameterization of the roots. (Kriegl, Losik, Michor [**KLM04**], 2004).
- If all a_i are of quasi-analytic class C^M, then there exists a parameterization of the roots of the same class C^M. (Chaumat, Chollet [**CC04**], 2004).
- For any given modulus of continuity ω there is a non-negative C^∞ function which does not admit a $C^{1,w}$ choice of its square roots. (Bony, Broglia, Colombini, Pernazza [**BBCP06**], 2005).

These results are best possible in their conclusions; the last statement shows, for instance, that one cannot expect more than twice differentiability for the roots. However, the conditions in some of the listed results may still be refined. In this concern an approach was recently presented in [**LR07**].

We shall give in this book a complete outline with complete proofs of the present standard of knowledge in this field of research.

The problem of choosing roots of regular curves of hyperbolic polynomials in a regular manner allows the following reformulation in terms of representation theory: Let the symmetric group S_n act in $\mathbb{R}^n$ by permuting the coordinates. The coordinates correspond to the roots of P. Consider the polynomial mapping $\sigma = (\sigma_1, \ldots, \sigma_n) : \mathbb{R}^n \to \mathbb{R}^n$ whose components σ_j are the elementary symmetric functions

$$\sigma_j(x_1, \ldots, x_n) = \sum_{1 \le i_1 < \cdots < i_j \le n} x_{i_1} \cdots x_{i_j}.$$

The σ_j represent the coefficients of P, via Vieta's formulas. Now the question of interest reformulates to: Given a regular curve $P : \mathbb{R} \to \sigma(\mathbb{R}^n) \subseteq \mathbb{R}^n$, where regularity of P refers to viewing it as curve in $\mathbb{R}^n$, is it possible to find a regular lift x of P, i.e., a regular curve $x : \mathbb{R} \to \mathbb{R}^n$ satisfying $\sigma \circ x = P$?

There is the following natural generalization of the above problem: Let $\rho : G \to \mathrm{O}(V)$ be an orthogonal representation of a compact Lie group G on a real finite dimensional Euclidean vector space V. It is well known that the algebra $\mathbb{R}[V]^G$ of invariant polynomials on V is finitely generated. Let $\sigma_1, \ldots, \sigma_n$ be a system of such generators, without loss all homogeneous of positive degree. Then the mapping $\sigma = (\sigma_1, \ldots, \sigma_n) : V \to \mathbb{R}^n$ is proper, separates orbits, and thus induces a homeomorphism between its image $\sigma(V)$ and the orbit space V/G. Hence we may identify the orbit space V/G with the semialgebraic subset $\sigma(V) \subseteq \mathbb{R}^n$. Suppose we are given a regular curve $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ in the orbit space, where regularity of c refers to viewing it as curve in $\mathbb{R}^n$. Does there exist a regular lift $\bar{c}$ of c, i.e., a regular curve $\bar{c} : \mathbb{R} \to V$ satisfying $\sigma \circ \bar{c} = c$?

One motivation for treating this general question roots in the study of smooth structures on orbit spaces. A set X can be given a notion of smoothness by declaring the family of smooth curves in X and the family of smooth functions on X. Such a set with smooth structure is called a Frölicher space after A. Frölicher who studied this notion in [**Frö80**] and [**Frö81**]. Note that this notion of smoothness generalizes the ordinary notion of smoothness on smooth manifolds. The orbit space $V/G = \sigma(V) \subseteq \mathbb{R}^n$ of a compact Lie group representation can be endowed with a Frölicher structure by declaring a curve c in $\sigma(V)$ smooth if $\iota \circ c$ is smooth, where $\iota : \sigma(V) \to \mathbb{R}^n$ denotes the inclusion map. The smooth functions on $V/G = \sigma(V)$ are then those whose composition with the smooth curves is smooth. It follows that the mapping $\sigma : V \to \sigma(V)$ which can be identified with the orbit projection is smooth, i.e., for any smooth curve in V its composition with σ is a smooth curve in $\sigma(V)$. The question posed in the previous paragraph amounts to the following: Are the structure curves of a quotient of a Frölicher space liftable as structure curves?

By the results we already know for the special case of the standard representation of the symmetric group S_n in $\mathbb{R}^n$, the answer to that question is no.

We dedicate the second part of this book to the detailed investigation of that generalized problem. Most results for the special case of perturbations of hyperbolic polynomials have been carried over to the general situation recently:

- Real analytic curves in the orbit space admit real analytic lifts, locally. Smooth (i.e. C^∞) curves in the orbit space allow global smooth lifts, if roughly speaking the curve does not meet lower dimensional strata of the orbit space with infinite order of flatness. If the representation ρ is polar, i.e., there exist linear subspaces in V meeting all orbits orthogonally, then the lifts may be chosen orthogonal to the orbits and unique up to a fixed transformation from G. (Alekseevsky, Kriegl, Losik, Michor [**AKLM00**], 2000).

- If the curve in the orbit space is of class C^d, where d is the maximal degree in a minimal system of homogeneous generators of $\mathbb{R}[V]^G$, then there exists a differentiable lift to the representation space. (Kriegl, Losik, Michor, R. [**KLMR05**], 2005).
- Let G be finite and write $V = V_1 \oplus \cdots \oplus V_l$ as orthogonal direct sum of irreducible subspaces V_i. Let k_i be the minimal cardinality of non-zero orbits in V_i and put $k = \max\{d, k_1, \ldots, k_l\}$. For a curve c in the orbit space we have:
  - If c is C^k, then there exists a differentiable lift of c with locally bounded derivative.
  - If c is C^{k+d}, then any differentiable lift of c is actually of class C^1.
  - If c is C^{k+2d}, then there exists a twice differentiable lift of c.

  (Kriegl, Losik, Michor, R. [**KLMR06**], 2005).

The examples for hyperbolic polynomials show that also these results are best possible in their conclusions.

The results on the abstract lifting problem may be applied to the study of perturbations of hyperbolic polynomials again. In this way one can obtain refinements of the results in the case that the polynomials have certain additional symmetries. More precisely, assume that the roots $x_1(t), \ldots, x_n(t)$ of the curve of polynomials $P(t)$ fulfill some linear relations, i.e., there is a linear subspace U of $\mathbb{R}^n$ such that $(x_1(t), \ldots, x_n(t)) \in U$ for all t. Then the curve $P(t)$ lies in the semialgebraic subset $\sigma(U) \subseteq \sigma(\mathbb{R}^n) = \mathbb{R}^n/\mathrm{S}_n$. The symmetries of the roots of $P(t)$ are represented by the action of W on U which is inherited from the S_n-action on $\mathbb{R}^n$, namely, $W = N(U)/Z(U)$ where $N(U)$ is the subgroup of elements in S_n preserving U and $Z(U)$ is the subgroup of elements in S_n leaving U pointwise fixed. If the restrictions $\sigma_i|_U$, $1 \le i \le n$, of the elementary symmetric functions σ_i to U generate the algebra $\mathbb{R}[U]^W$, then we may consider $P(t)$ as a curve in the orbit space $U/W = \sigma(U)$. Regular lifts over $\sigma|_U$ to U provide then regular parameterizations of the roots of $P(t)$. Since $P(t)$ lies in a proper subset of $\mathbb{R}^n/\mathrm{S}_n = \sigma(\mathbb{R}^n)$, the conditions for regular lifting are in general weaker. This idea is developed in [**LR07**].

As indicated we dispose now of a quite complete picture of the perturbation theory of hyperbolic polynomials and the related problem of lifting curves over invariants. The interesting problems apart from a few seem to be solved. Let us list a few remaining open questions worthwhile investigating:

- The quadratic case, where for a non-negative function it suffices to be C^2 for the existence of C^1 square roots of that function and C^4 for the existence of twice differentiable square roots, indicates that the regularity conditions in the statements above are not weakest possible.
- If no restrictions on the polynomials P are imposed and the roots have in general a non-zero imaginary part, very little can be expected. Simplest examples as $x^2 - t = 0$, $t \in \mathbb{R}$, show that even for a real analytic curve of polynomials the roots cannot in general satisfy a Lipschitz condition. Still the roots may possess a weaker regularity property; absolute continuity. In fact, it was shown in [**Spa99**] that sufficiently differentiable curves of monic polynomials of degree 2 and 3 allow absolutely continuous parameterizations of its roots. For arbitrary degree the problem is open as far as we know.
- The results on the existence of differentiable lifts with locally bounded derivatives, C^1 lifts, and twice differentiable lifts in the last point of the previous list can easily be carried over to polar representations. For general representations, of which polar representation constitute only a minor part, it is unclear whether similar statements are true and how to prove or disprove them.

Of course we do not claim that this list is complete.

This monograph is intended as a broad and detailed systematic presentation of the study of perturbations of hyperbolic polynomials and the generalized problem of lifting curves over invariants which comprises classical theorems as well as most resent results. One of its principal items is a rigorous and understandable exposition of Bronshtein's important theorem in [**Bro79**] which is up to now not available in the mathematical literature. The second part develops rather extensively the prerequisites and tools for the treatment of the generalized lifting problem. The third part is dedicated to a few related problems which either are applications or are proximate questions of the methods and results presented in the first two parts. We think that this book can be a useful presentation of a field of ongoing mathematical research. It combines a bunch of important results which up to now were scattered in a multitude of pieces. The monograph is self-contained with only few exceptions. Those results which are presented without proof are equipped with references to the bibliography, and mostly they are important only marginally for the development and goal of the work. We are aware of the fact that this monograph does not provide an all-embracing presentation, in particular, of the rich theory of hyperbolic polynomials which still is advancing in different directions. Let us summarize the contents of this book.

In chapter 1 we explain two motivations for studying perturbations of hyperbolic polynomials and liftings of smooth curves over invariants. The first one roots in the theory of partial differential equations. We try to explain where and how Bronshtein [**Bro80**] uses his theorem on the existence of differentiable roots with locally bounded derivative of a smooth curve of hyperbolic polynomials in order to show well-posedness of the hyperbolic Cauchy problem with non-constant coefficients in Gevrey space. The second motivation is inspired by the study of smooth structures on singular spaces. We introduce the notion Frölicher space, indicate how orbit spaces of finite dimensional orthogonal representations of compact Lie groups may be considered as Frölicher spaces, and point out how the problem of lifting curves over invariants smoothly fits into that concept.

Chapter 2 collects results on the continuity of the roots of polynomials depending on their coefficients. Initially we do not impose any restrictions on the polynomials, so its roots are complex numbers. First an elementary approach leads to the statement that the roots of a polynomial of degree n, as a function of the coefficients, satisfy a Hölder condition of order $\frac{1}{n}$, see section 2.1. Next an application of Rouché's theorem yields the continuity of the roots of a holomorphic equation depending continuously on a parameter, see section 2.2. In these results continuity is understood as continuity of the set of roots as a whole. A different question is whether it is possible to find single-valued continuous functions which parameterize the roots of polynomials. Such a parameterization is possible if either the parameter the coefficients of a polynomial depend on is real or the polynomial is always hyperbolic, see section 2.3.

Chapter 3 is dedicated to the study of the space of hyperbolic monic polynomials of a fixed degree. We introduce the Bezoutiant, a particular matrix associated to a polynomial, and prove Sylvester's version of Sturm's theorem, in a modern approach due to C. Procesi [**Pro78**], which says that a polynomial is hyperbolic if and only if its Bezoutiant is positive semidefinite, see section 3.1. That characterization provides an explicit set of inequalities defining the space of hyperbolic polynomials of degree n when identified with a semialgebraic subset in $\mathbb{R}^n$ via its coefficients. The rest of the chapter deals with the study of the geometric properties of the space of hyperbolic polynomials due to V.I. Arnol'd [**Arn86**], A.B. Givental' [**Giv87**], V.I. Kostov [**Kos89**], and I. Meguerditchian [**Meg92**]. For instance, we investigate the 'escape' from the space of hyperbolic polynomials in terms of the multiplicity vector, i.e., the tuple of multiplicities of the ordered roots of a polynomial, see section 3.3. Moreover, we show in section 3.4 that the space of hyperbolic polynomials of a fixed degree and vanishing first and bounded second coefficient is Whitney regular, that means that any two points can be connected by a piecewise smooth curve within that space whose length is dominated by the Euclidean distance of the points times an independent constant.

In chapter 4 we follow the approach of Alekseevsky, Kriegl, Losik, Michor [**AKLM98**] to the problem of finding smooth parameterizations of the roots of hyperbolic polynomials. First the quadratic case is discussed, see section 4.1. Already this simple special case shows that the problem is very delicate: Even starting with a non-negative C^∞ function, one cannot expect more than twice differentiability for its square roots. Next the general problem for arbitrary degree is tackled. In section 4.2 an algorithm is constructed which factors a smooth curve of hyperbolic polynomials in a smoothly solvable and a potentially smoothly unsolvable part. In section 4.3 global smooth parameterizations which are unique up to permutations of the roots of a smooth curve of hyperbolic polynomials are constructed under the assumption that no two roots meet of infinite order of flatness. Moreover, it is proved that global differentiable choices of the roots exist whenever the hyperbolic polynomials of degree n depend in a C^n way on a real parameter. In section 4.4 a simple proof for the classical Rellich theorem, stating that real analytic curves of hyperbolic polynomials allow real analytic roots, is given. We show in section 4.5 that a quasi-analytic curve of hyperbolic polynomials allows a quasi-analytic parameterization of its roots, which is due to Chaumat and Chollet [**CC04**]. The chapter ends with some glance to the situation when hyperbolicity is absent, see section 4.6.

Chapter 5 presents a detailed elaboration of the sketchy paper [**Bro79**] due to Bronshtein. We start with a discussion of Bronshein's theorem when the polynomials are of degree three. This shortens and simplifies the general proof essentially, but uses the whole machinery of argumentation. That should help the reader to get an idea of Bronshtein's method which is based on the classical Sturm algorithm, since the general proof is rather long, involved, and technical. Next we present Bronshtein's proof of the fact that a C^n curve of hyperbolic polynomials of degree n admits differentiable roots. The advantage of Bronshtein's approach with respect to the approach of Alekseevsky, Kriegl, Losik, Michor at this point is that Bronshtein is able to derive a polynomial equation for the derivatives of the roots which accounts for its dependence on the parameter, see section 5.3. This will be of decisive importance for proving local boundedness of the derivative of the roots. Section 5.4 provides the necessary preliminaries collecting some simple facts on the coefficients and the roots of hyperbolic polynomials and their interrelation. Finally in section 5.5 we prove local boundedness of the derivative of the roots of a C^n curve of hyperbolic polynomials of degree n. Basically, it is shown that the coefficients of the equation the derivatives of the roots have to satisfy are locally bounded. Finally, we shall see that this one dimensional result implies that in the case that the coefficients of a hyperbolic polynomial of degree n depend in a C^n way on a multi-dimensional parameter, then there is a continuous parameterization of its roots which is locally Lipschitz.

In chapter 6 we follow Wakabayashi's approach [**Wak86**] and prove his more general version of Bronshtein's theorem. Wakabayashi's method is more conceptual. A main ingredient is the splitting operator $P \mapsto P + sP'$ where P is a hyperbolic polynomial and $s \in \mathbb{R}$. This operator preserves hyperbolicity and reduces the multiplicity of the roots of P. Moreover, one can give an estimate for the deviation the roots are subjected to under this operator. In section 6.2 we introduce the notion 'microhyperbolicity' and discuss some properties. This is based on [**Hör63**] and [**Hör83a**]. This notion is very similar to the notion of hyperbolic polynomials mentioned at the beginning: A real analytic function F on an open set $U \subseteq \mathbb{R}^n$ is called microhyperbolic with respect to $\Theta \in \mathbb{R}^n$ if there is a positive continuous function $x \mapsto t(x)$ such that $F(x + it\Theta) \neq 0$ for $0 < t < t(x)$ and $x \in U$. Microhyperbolicity is used in section 6.3 where we prove Wakabayashi's theorem which (basically) states that the increasingly ordered roots of a $C^{k,\alpha}$ $(0 < \alpha \leq 1)$ curve of hyperbolic polynomials satisfy a Hölder condition with exponent $\min\{1, \frac{k+\alpha}{n}\}$ on any open bounded interval. Bronshtein's theorem follows from that statement.

We prove in chapter 7 that for a curve P of hyperbolic polynomials of degree n any differentiable parameterization of the roots is actually C^1 if P is C^{2n}, and that there exists a twice differentiable parameterization of the roots if P is C^{3n}. The first statement is due to Mandai [**Man85**] and the second to Kriegl, Losik, Michor [**KLM04**]. Note that the conclusions are best possible as already ascertained above. The proof of those results relies on Bronshtein's boundedness result which implies that the statements hold at any point of the parameter domain. The rest of the proof carried out by induction on the degree n shows that these local parameterizations may be glued in order to get global C^1 and twice differentiable parameterizations of the roots, respectively.

We begin part 2 of this book with chapter 8 which is dedicated to the reformulation and generalization of the problem of choosing the roots of hyperbolic polynomials in a regular way in terms of a lifting problem from the orbit space V/G of a finite dimensional orthogonal compact Lie group representation $G \to \mathrm{O}(V)$ to the representation space V. We discuss the general setting and prove that the mapping $\sigma = (\sigma_1, \ldots, \sigma_n) : V \to \mathbb{R}^n$, where $\sigma_1, \ldots, \sigma_n$ constitutes a system of homogeneous generators of $\mathbb{R}[V]^G$, is proper, separates orbits, and induces a homeomorphism between V/G and $\sigma(V)$. Then we introduce a functional structure on V/G, by declaring a function on V/G smooth if its composition with the orbit projection is smooth and state Schwarz's theorem which provides a characterization of the G-invariant smooth functions on V, namely, it says that $\sigma^* : C^\infty(\mathbb{R}^n) \to C^\infty(V)^G$ is surjective. Finally, we show that the problem of lifting curves over invariants is independent of the choice of the generators $\sigma_1, \ldots, \sigma_n$.

Chapter 9 provides the necessary background from the theory of isometric actions of Lie groups on manifolds required for the study of the mentioned lifting problem. We try to give a accurate and complete presentation of the results which are relevant for our purpose and at the same time we are keen not to depart to much from our subject, the lifting problem. In section 6.1 we prove Schwarz's theorem. We do not follow Schwarz's original proof in [**Sch75**], but we prove a stronger variant with simpler proof due to Mather [**Mat77**]. Mather could show that $\sigma^* : C^\infty(\mathbb{R}^n) \to C^\infty(V)^G$ is even split surjective, i.e., a linear continuous right inverse exists. Next we define the generalized Bezoutiant, namely, $B(v) = (\langle d\sigma_i(v) \mid d\sigma_j(v)\rangle)_{ij}$, $(v \in V)$, which reduces to the Bezoutiant associated to polynomials if as generators for $\mathbb{R}[\mathbb{R}^n]^{\mathrm{S}_n}$ are chosen $\sigma_k = \frac{1}{k} s_k$ where s_k denotes the k-th Newton polynomial. As for polynomials we find a description of the image $\sigma(V)$ by finitely many equations and inequalities due to Procesi and Schwarz [**PS85**]. In fact, $\sigma(V) = \{z \in V/\!\!/G : \tilde{B}(z) \text{ positive semidefinite}\}$, where $\tilde{B}$ is uniquely given by $B = \tilde{B} \circ \sigma$ and $V/\!\!/G = \{y \in \mathbb{R}^n : P(y) = 0 \text{ for all } P \in I\}$ with I the ideal of all relations between $\sigma_1, \ldots, \sigma_n$. We prove that theorem in section 9.3. To this end we need some facts on linear reductive groups, we present Luna's slice theorem, study complexifications of compact Lie groups, and apply results on stability due to Kempf and Ness [**KN79**] and Dadok and Kac [**DK85**]. In section 9.4 we discuss orbit types and slices of G-manifolds. We prove the differentiable slice theorem, a most useful tool for the further treatment of our problem. Then, proper actions are introduced and investigated. In section 9.5 we point out that in the setting of our lifting problem the existence of slices is guaranteed and we describe a reduction procedure of the lifting problem: By the slice theorem, we may reduce the local problem to the slice representation $G_v \to \mathrm{O}(N_v)$, where G_v is the isotropy group of v and N_v is the orthogonal complement of the tangent space at v to the orbit through v. Finally, we give a description of the space $N_v^{G_v}$ of G_v-invariant points in N_v in terms of $\sigma_1, \dots, \sigma_n$. We conclude the chapter with the description of the stratification of the representation space V and the orbit space V/G by orbit types in section 9.6. We prove that the partition into connected components of subsets with constant orbit type indeed provides a stratification in the sense of Mather [**Mat73**] and satisfies Whitney's condition (b), see also [**Pfl01**]. Moreover, we show that the orbit type stratification of V/G when identified with $\sigma(V)$ coincides with its (primary) stratification as semialgebraic set, which is due to Bierstone [**Bie75**].

We lift curves over invariants real analytically and smoothly in chapter 10. In analogy to the polynomial case considered in chapter 4 we construct an algorithm which solves the lifting problem locally. First the real analytic case is considered and this case indicates which additional conditions have to be imposed in order to be able to lift smooth curves smoothly. For this purpose the notion of 'normal nonflatness' is introduced: Roughly speaking a curve in the orbit space is normally nonflat at some point if at that point it does not meet a lower dimensional stratum with infinite order of flatness. The problem of lifting smooth curves smoothly can be solved locally wherever the curve is normally nonflat. If the curve is normally nonflat everywhere, we can construct a global smooth lift due to the fact that there are smooth partitions of unity, see section 10.2. In section 10.3 we introduce polar representations, i.e., there exists a linear subspace $\Sigma \subseteq V$ called section which meets each orbit orthogonally. The section Σ inherits the action of the generalized Weyl group $W(\Sigma)$ given as the quotient of the subgroup of elements in G which preserve Σ by the subgroup of elements which fix Σ pointwise. The generalized Weyl group is finite, and, by the generalization of Chevalley's restriction theorem [**DK85**] and [**Ter85**] the algebras $\mathbb{R}[V]^G$ and $\mathbb{R}[\Sigma]^{W(\Sigma)}$ are isomorphic by restriction, whence $V/G = \Sigma(V)$ and $\Sigma/W(\Sigma) = \sigma|_\Sigma(\Sigma)$ are isomorphic as stratified spaces. For polar representations we show that real analytic curves in the orbit space allow global orthogonal (to the orbits) real analytic lifts and smooth everwhere normally nonflat curves in the orbit space allow global smooth lifts. Moreover the lifts are unique up to a constant transformation of G. This chapter is based on [**AKLM00**].

In chapter 11 we tackle the same lifting problem under weaker differentiability conditions. More precisely, we do not impose any restriction on our regular curves in the orbit space and observe what we can still achieve. We show first that continuous curves admit continuous lifts. This fact is true as well in a more general setting: Montgomery and Yang [**MY57**] proved that continuous curves in the orbit space of a Hausdorff topological space endowed with a continuous action by a compact Lie group have continuous lifts. In section 11.2 we associate to the representation $\rho : G \to \mathrm{O}(V)$ the integer $d = d(\rho)$, namely, the maximal degree in a minimal system of homogeneous generators of $\mathbb{R}[V]^G$. Then $d(\rho)$ is well-defined and independent of the choice of generators, and, for any slice representation ρ' of ρ, we have $d(\rho') \leq d(\rho)$. Finally, following [**KLMR05**], we show in section 11.4 that any C^d curve in the orbit space V/G has a global differentiable lift to V. Not that examples show that here the assumption cannot be weakened.

We are able to obtain much stronger results if we restrict to polar representations. In chapter 12, which is based on [**KLMR06**], we first define what it means for a representation to have property $(\mathcal{B}_k)$ which imitates Bronshtein's result: There exists a neighborhood U of 0 in $V/G = \sigma(V)$ such that each C^k curve in U admits a local differentiable lift to V with locally bounded derivative. We study this notion for representations of finite groups. Then we show in dependence on chapter 7 that any differentiable lift of a C^{k+d} curve in the orbit space of a representation ρ of a finite group with property $(\mathcal{B}_k)$ is actually C^1, and any C^{k+2d} curve admits a twice differentiable lift. Obviously, these results generalize to polar representations, since we may reduce to the representation of the finite generalized Weyl group $W(\Sigma)$ on a section Σ. Examples show that the conclusions in these results are best possible. It turns out, see section 12.6, that all representations ρ of finite groups, and hence all polar representations, have property $(\mathcal{B}_k)$, where k is specified as follows: Write $V = V_1 \oplus \cdots \oplus V_l$ as orthogonal direct sum of irreducible subspaces V_i. Let k_i be the minimal cardinality of non-zero orbits in V_i and put $k = \max\{d, k_1, \ldots, k_l\}$. We discuss in detail the cases when G is a finite reflection group or a finite rotation group in two and three dimensions. Finally, we show in section 12.9 that any C^k mapping $f : \mathbb{R}^m \to V/G$ into the orbit space of a polar representation allows a continuous lift $\bar{f} : \mathbb{R}^m \to V$ which is Lipschitz locally.

As an application of the abstract lifting problem we prove some refinements for the perturbation problem of hyperbolic polynomials in chapter 13, based on [**LR07**]. The idea is the following: Let us view the space Hyp_n of hyperbolic polynomials of degree n as semialgebraic subset of $\mathbb{R}^n$ and consider a semialgebraic subset $X \subseteq \mathrm{Hyp}_n$. A smooth curve c in X may be viewed as a smooth curve of hyperbolic polynomials, which in general must fulfill weaker genericity conditions in order to be smoothly liftable than in the case it would not be restricted to X. First we provide a class of cases where this strategy is applicable: By a result due to Smith and Strong [**SS87**], the orbit space of any faithful representation of a finite group can be embedded in a finite product $\mathrm{Hyp}_{n_1} \times \cdots \times \mathrm{Hyp}_{n_l}$, see section 13.2. Then we change the point of view. We suppose that a smooth curve of hyperbolic polynomials $P(t)$ satisfies certain symmetries, more precisely, there is a linear subspace U of $\mathbb{R}^n$ such that $(x_1(t), \ldots, x_n(t)) \in U$ for all t, where $x_1(t), \ldots, x_n(t)$ are the roots of $P(t)$. The symmetries of $P(t)$ are represented by the action of $W(U)$ on U which is inherited from the action of symmetric group S_n on $\mathbb{R}^n$ in the same way as the generalized Weyl group above. If, additionally, the restrictions to U of the elementary symmetric functions generate $\mathbb{R}[U]^{W(U)}$, the orbit space $U/W(U)$ is naturally embedded in Hyp_n. Then the curve $P(t)$ lies in $U/W(U)$ and can be lifted to U, by applying the abstract lifting procedure. Evidently, such a lift represents a parameterization of the roots of $P(t)$. We show in section 13.3 that normal nonflatness as curve in Hyp_n implies normal nonflatness as curve in $U/W(U)$. Consequently, the depicted method in general provides improvements for the conditions of smooth lifting. Furthermore, it leads to refinements also for finite differentiability, see section 13.5. At the end of the chapter a class of examples is constructed and the construction is applied to finite reflection groups, see section 13.6 and section 13.7.

Part 2 of this treatise ends with an appendix in which a complete list of basic invariants for finite irreducible reflection groups and finite rotation groups in two and three dimensions is given.

In part 3 we treat some related topics. We begin in chapter 14 with the perturbation theory of linear operators. Section 14.1 deals with the problem of parameterizing the eigenvalues and the eigenvectors of matrices smoothly. First we recall some classical results due to Rellich [**Rel37**], [**Rel42**], [**Rel69**], and Kato [**Kat76**]: The eigenvalues, eigenprojections, and eigennilpotents of a holomorphic curve of $(n \times n)$-matrices are holomorphic with at most algebraic singularities at discrete points. The eigenvalues and eigenvectors of a real analytic curve of Hermitian matrices may be parameterized real analytically. The eigenvalues of a C^1 curve of symmetric matrices can be chosen C^1, globally. Then, as an application of the results obtained in chapter 4, we show that the eigenvalues and the eigenvectors of a smooth curve of Hermitian matrices can be parameterized smoothly, globally, if no two of the continuously parameterized eigenvalues meet of infinite order of flatness. This is based on [**AKLM98**]. In section 14.2 we discuss the corresponding perturbation problem for unbounded self-adjoint compact operators with compact resolvent. The consideration of the real analytic case is again due to Rellich [**Rel42**]. We follow [**AKLM98**] and [**KM03**] and prove that for a curve $t \mapsto A(t)$ of unbounded self-adjoint operators in a Hilbert space with common domain of definition and with compact resolvent the following is true: If $A(t)$ is smooth in t and no two of the continuously parameterized eigenvalues meet of infinite order, then the eigenvalues and the eigenvectors can be parameterized smoothly, globally. If $A(t)$ is smooth in t, then the eigenvalues may be globally parameterized twice differentiably. If $A(t)$ is $C^{1,\alpha}$ for some $\alpha > 0$, then the eigenvalues may be globally parameterized in a C^1 way.

We come back to the notion of hyperbolic polynomials, introduced at the very beginning of this introduction, in chapter 15. We shall give an idea of the connection of this notion originated from the theory of partial differential equations to convex analysis as pointed out in [**BGLS01**]. If a homogeneous polynomial P of degree m is hyperbolic with respect to N, then for any ξ we may write $P(\xi + \tau N) = P(N)\prod_{i=1}^{m}(\tau + \lambda_i(\xi))$ and assume that $\lambda_1(\xi) \geq \lambda_2(\xi) \geq \cdots \geq \lambda_m(\xi)$. Gårding [**Går51**], [**Går59**] showed that the largest root λ_1 is a sublinear function, i.e., positively homogeneous and convex. We indicate how to construct new hyperbolic polynomials and discuss a few features. Finally we state a theorem due to Nuij [**Nui68**] which deals with the topological properties of the space of polynomials hyperbolic with respect to some given vector.

In chapter 16 we address related lifting problems. In section 16.1 we follow [**Los01**] and study the lifting of diffeomorphisms of orbit spaces. The orbit space V/G of a compact Lie group representation has a real analytic or smooth structure defined by the sheaf of invariant functions, as mentioned above. Let $f : V/G \to V/G$ be a diffeomorphism in the usual categorical meaning. A diffeomorphism $F : V \to V$ is called a lift of f if $\pi \circ F = f \circ \pi$, where $\pi : V \to V/G$ is the orbit projection. It is shown that for finite G any smooth (real analytic) diffeomorphism of V/G has a smooth (real analytic) lift to V. Any two such lifts are unique up to a fixed transformation from G. Moreover, the component group $\pi_0(\mathrm{Diff}(V/G))$ of the group of diffeomorphisms $\mathrm{Diff}(V/G)$ is studied. In section 16.2 we consider representations of finite groups G in a complex vector space V. Then the orbit space V/G coincides with the categorical quotient $V/\!\!/G$ which is a normal affine variety. Thus, the orbit space V/G has the natural structure of a complex analytic set and there are several types of morphisms into V/G. We investigate the conditions for lifts of germs of holomorphic morphisms at 0 from $\mathbb{C}^p$ to V/G, for lifts of regular maps from $\mathbb{C}^p$ to V/G, and for lifts of formal morphisms from $\mathbb{C}^p$ to V/G, i.e., the morphisms of the $\mathbb{C}$-algebra $\mathbb{C}[V/G]$ to the ring of formal power series $\mathbb{C}[[X_1, \ldots, X_p]]$ in variables $X_1, \ldots, X_p$. This section is based on [**KLMR08**].

I would like to thank Peter W. Michor for encouraging me to write this monograph. Moreover, I am indebted to him as well as Andreas Cap, Stefan Haller, Andreas Kriegl, and Mark Losik for all the constructive discussions.

Contents

CHAPTER 1

Motivation

1.1. Hyperbolic polynomials in the theory of PDEs

1.1.1. The Cauchy problem. Let us consider the equation

$$P(x,D)u(x) \equiv \sum_{|\alpha|\le m} a_\alpha(x) D^\alpha u(x) = f(x), \tag{1.1}$$

where $D_{x_j} = -i\frac{\partial}{\partial x_j}$, $\alpha = (\alpha_0, \alpha') = (\alpha_0, \alpha_1, \dots, \alpha_l) \in \mathbb{N}^{1+l}$, $x = (x_0, x') \in \Omega \subseteq \mathbb{R}^{1+l}$, and Ω is a neighborhood of the origin. The plane $x_0 = \text{const}$ is supposed to be non-characteristic, i.e., $a_{(m,0,\dots,0)}(x) \ne 0$ for all $x \in \Omega$. For equation (1.1) we pose the *Cauchy problem* (or *initial value problem*)

$$D^j_{x_0} u(x)\big|_{x_0=0} = g_j(x'), \quad 0 \le j \le m-1. \tag{1.2}$$

We are interested in conditions on the operator P and the functions f and g_j which guarantee the existence of a solution of the problem (1.1) & (1.2) in some neighborhood $\Omega' \subseteq \Omega$ of the origin.

The classical Cauchy-Kowalewski theorem, [**Hör83a**, Theorem 9.4.5], provides an example of such conditions, namely, sufficient analyticity of all the given functions. Here the domain of existence of the solution depends both on the operator P and on the radius of analyticity of f and the g_j. If the operator P is *hyperbolic* in Ω relative to x_0, i.e., all roots $\lambda_j(x,\xi')$ $(1 \le j \le m)$ in ξ_0 of the principal symbol $P_m(x,\xi_0,\xi') = \sum_{|\alpha|=m} a_\alpha(x)\xi^\alpha$ are real for any $(x,\xi') \in \Omega \times \mathbb{R}^l$, then, as was shown by Bony and Schapira [**BS73b**],[**BS73a**], the domain of existence depends only on P. Komatsu [**Kom77**] has proved the converse: If there is a neighborhood Ω of the origin, not depending on the radius of analyticity of f and the g_j, in which there exists a classical solution (or even a hyperfunction solution), then P is hyperbolic in Ω relative to x_0. It is known that also a replacement of the space of analytic functions in the Cauchy-Kowalewski theorem by spaces of lower smoothness requires P to be hyperbolic.

It is well-known that if all the characteristic roots $\lambda_j(x,\xi')$ $(\xi' \ne 0)$ of the principal symbol $P_m(x,\xi)$ of a hyperbolic operator P are distinct, then the Cauchy problem (1.1) & (1.2) is well-posed in the space of C^∞ functions. Recall that the Cauchy problem is said to be *well-posed* if there exists a unique solution depending continuously on the initial data.

In the presence of multiple roots this is not the case, and to characterize conditions for solvability, one needs to recourse to Gevrey spaces.

1.1.2. Definition. We denote by $G^s(\Omega)$ the set of all $f \in C^\infty(\Omega)$ such that for every compact set $K \subseteq \Omega$ there is a constant C_K with

$$|D^\alpha f(x)| \le C_K (C_K |\alpha|^s)^{|\alpha|}, \quad x \in K,$$

for all multi-indices α. The class $G^s(\Omega)$ is called *Gevrey class of order s*. It is a ring which is closed under differentiation.

1.1.3. In [**Bro80**] M.D. Bronshtein proves that if all the given functions belong to the Gevrey space $G^s(\Omega)$ with $s = \frac{r}{r-1}$ then the Cauchy problem (1.1) & (1.2) is solvable for an arbitrary hyperbolic operator the multiplicity of whose characteristic roots does not exceed r. It is in general not possible to go beyond the limits of Gevrey spaces of order $\frac{r}{r-1}$. Here is the precise statement.

THEOREM. *Suppose that the multiplicity of the roots $\lambda_j(x,\xi')$ of the highest symbol $P_m(x,\xi)$ of a hyperbolic operator P does not exceed $r > 1$ for all $(x,\xi') \in \Omega \times \mathbb{R}^l\backslash\{0\}$. Suppose also that all given functions a_α, f, and g_j of the Cauchy problem* (1.1) *&* (1.2) *belong to the space $G^s(\Omega)$ with $1 \le s \le \frac{r}{r-1}$. Then in some neighborhood $\Omega' \subseteq \Omega$ of the origin (which depends on P, f, and the g_j) there exists a solution of the Cauchy problem* (1.1) *&* (1.2) *that belongs to $G^s(\Omega')$ and is unique in the space $\mathscr{D}'(\Omega')$ of generalized functions.*

Note that it is also possible to formulate the theorem for finite differentiability as well as for a system of equations; see [**Bro80**]. When dealing with finite differentiability, one has to replace the spaces $G^s(\Omega)$ by $G^s_{C^k}(\Omega)$ $(k > 3(m+l+2))$ in the conditions of the above theorem. Here $G^s_{C^k}(\Omega)$ denotes the set of functions f having continuous derivatives $D^{\alpha_0}_{x_0} D^{\alpha'}_{x'} f$ in Ω for any $\alpha_0 \le k$ and $\alpha' \in \mathbb{N}^l$, and satisfying on each compact $K \subseteq \Omega$ the estimate

$$|D^{\alpha_0}_{x_0} D^{\alpha'}_{x'} f(x)| \le C_K^{|\alpha|+1} |\alpha'|^{s|\alpha'|}$$

for all $\alpha' \in \mathbb{N}^l$, $0 \le \alpha_0 \le k$, and for all $x \in K$.

1.1.4. Bronshtein's proof of theorem 1.1.3 is by constructing a parametrix, i.e., an operator giving a solution of the Cauchy problem to within an integral operator in certain L^2-spaces with a weight.

Let us try to give an idea where in the proof of theorem 1.1.3 the local boundedness of the derivative of the roots of a hyperbolic polynomial depending smoothly on a real parameter, established by Bronshtein in [**Bro79**], is needed. The precise statement is the following (see theorem 5.5.13 and also theorem 6.3.1): *Suppose that the polynomial*

$$P(t,y)(x) = x^m + a_1(t,y)x^{m-1} + \cdots + a_m(t,y)$$

is hyperbolic, i.e., has only real roots, for any $(t,y) \in (-1,1) \times \mathcal{M}$, where $\mathcal{M}$ is a compact Hausdorff topological space, and the multiplicity of its roots does not exceed r. Furthermore, suppose that all partial derivatives $\frac{\partial^i}{\partial t^i} a_j(t,y)$ $(i = 0,\ldots,r; j = 1,\ldots,m)$ are continuous functions on $(-1,1) \times \mathcal{M}$. Then, for any compact subset $K \subseteq (-1,1) \times \mathcal{M}$, there exists a positive constant C_K such that, for all (differentiably chosen) roots $x_j(t,y)$ $(j = 1,\ldots,m)$ of P, we have the following estimate

$$\left|\frac{\partial}{\partial t} x_j(t,y)\right| < C_K \quad \textit{for all } (t,y) \in K.$$

The parametrix is constructed in the same way as for elliptic operators, which is possible since one can go into a complex domain in ξ_0. Indeed, let us consider

$$\Lambda_{H,s} = \{\zeta = (\zeta_0,\xi') \in \mathbb{C} \times \mathbb{R}^l : |\mathrm{Im}(\zeta_0)| > H(1 + |\xi'|^{\frac{1}{s}})\},$$

and suppose that $1 \le s \le \frac{r}{r-1}$. We look for the symbol of the parametrix $Q_N(x,\zeta)$ in the form $Q_N = Q'_1 + \cdots + Q'_N$, and Q'_N will as usual be defined recursively such that

$$P \cdot Q := \sum_{|\gamma| \le m} \frac{1}{\gamma!} P^{(\gamma)}(x,\zeta) Q_{N(\gamma)}(x,\zeta) = 1 + P(x,\zeta) Q'_{N+1}(x,\zeta),$$

i.e., we set $Q = \frac{1}{P}$ and $Q'_{N+1} = \frac{1}{P}(1 + P \cdot Q_N)$. Here we use the notation $P^{(\alpha)}_{(\beta)} = D^\beta_x (\frac{\partial}{\partial \xi})^\alpha P$. We have

$$P \cdot Q_N = 1 + R_N,$$
$$Q_N(x,\zeta) = \frac{1}{P(x,\zeta)} \sum_{|\alpha|=|\beta|\le mN} A_{\alpha,\beta} \prod_j \frac{P^{(\alpha^j)}_{(\beta^j)}(x,\zeta)}{P(x,\zeta)},$$
$$R_N = \sum_{|\alpha|=|\beta|\le mN, q\ge N+1} \tilde{A}_{\alpha,\beta} \prod_{j=1}^q \frac{P^{(\alpha^j)}_{(\beta^j)}}{P},$$

where $|\alpha| = \sum_j |\alpha^j|$, $|\beta| = \sum_j |\beta^j|$, $\alpha^j, \beta^j \in \mathbb{N}^{1+l}$, and $A_{\alpha,\beta}, \tilde{A}_{\alpha,\beta} \in \mathbb{N}$. It is crucial for the proof to estimate the derivatives ${R_N}^{(\alpha)}_{(\beta)}$ of the remainder term, and hence of expressions of the type $\frac{P^{(\alpha)}_{(\beta)}}{P}$. To this end consider the following proposition.

Proposition ([**Bro80**, Proposition 3]). *Let*

$$P(x,\zeta) = \zeta_0^m + a_1(x,\xi')\zeta_0^{m-1} + \cdots + a_m(x,\xi') = \prod_j (\zeta_0 - \lambda_j(x,\xi'))$$

be a polynomial hyperbolic in ζ_0, and let $a_i(x,\xi') \in C^k(\Omega \times \mathbb{R}^l \backslash \{0\})$ be positively homogeneous functions of degree i. Then, for any compact $K \subseteq \Omega$, $0 \neq |\alpha + \beta| < k - m - \gamma_0$, $\gamma_0 \in \mathbb{N}$, and $H > 1$,

$$\left| \frac{P^{(\alpha_0+\gamma_0,\alpha')}_{(\beta)}(x,\zeta)}{P(x,\zeta)} \right| \le C_{K,k} \frac{(1+|\xi'|)^{|\beta| - \frac{(|\alpha|+|\beta|-1)}{s}}}{(H(1+|\xi'|)^{\frac{1}{s}} + |\zeta|^{\frac{1}{s}})^{\gamma_0+1}}$$

for all $(x,\zeta) \in K \times \Lambda_{H,s}$ and $|\xi'| > 1$.

Sketch of proof. We use induction on $|\alpha' + \beta|$. Suppose that $|\alpha' + \beta| = 0$. Then we have

$$\left| \frac{P^{(\alpha_0)}(x,\zeta)}{P(x,\zeta)} \right| = \left| \sum_{1 \le i_1 < \cdots < i_{\alpha_0} \le m} \prod_{j=1}^{\alpha_0} (\zeta_0 - \lambda_{i_j}(x,\xi'))^{-1} \right|$$
$$\le 2^m \left(\frac{16 C_K}{H(1+|\xi'|)^{\frac{1}{s}} + |\zeta|^{\frac{1}{s}}} \right)^{\alpha_0}$$

for all $(x,\zeta) \in K \times \Lambda_{H,s}$, since $|\zeta_0 - \lambda_j(x,\xi')| > (16C'_K)^{-1}\left(H(1+|\xi'|)^{\frac{1}{s}} + |\zeta|^{\frac{1}{s}}\right)$, where $C'_K = 1 + \max_{x\in K, j, \xi'} \left|\frac{\lambda_j(x,\xi')}{\xi'}\right|$; see [**Bro80**, Proposition 1].

Let $|\alpha'| + |\beta| = 1$. The roots λ_i have first order partial derivatives in the parameters (x,ξ') and

$$\left| \frac{\partial}{\partial \xi_j} \lambda_i(x,\xi') \right| < C_K, \quad \left| \frac{\partial}{\partial x_j} \lambda_i(x,\xi') \right| < C_K |\xi'|, \tag{1.3}$$

for all $(x,\xi') \in K \times \mathbb{R}^l \backslash \{0\}$ and $|\xi'| > 1$, by the fact mentioned above. Hence the proposition holds for $|\alpha'| + |\beta| = 1$:

$$\left| \frac{P^{(\alpha_0,\alpha')}(x,\zeta)}{P(x,\zeta)} \right| = \left| \sum \frac{\partial}{\partial \xi_j} \lambda_i(x,\xi') \prod_{q=1}^{\alpha_0+1} (\zeta_0 - \lambda_{i_q}(x,\xi'))^{-1} \right|$$
$$\le C''_K \left(H(1+|\xi'|)^{\frac{1}{s}} + |\zeta|^{\frac{1}{s}} \right)^{-(\alpha_0+1)} \quad (|\alpha'| = 1)$$

and, similarly,

$$\left|\frac{P^{(\alpha_0)}_{(\beta)}(x,\zeta)}{P(x,\zeta)}\right| \le C''_K|\xi'|\left(H(1+|\xi'|)^{\frac{1}{s}}+|\zeta|^{\frac{1}{s}}\right)^{-(\alpha_0+1)} \qquad (|\beta|=1).$$

Let us assume that the proposition holds for $|\alpha'|+|\beta|<v$. Consider $|\alpha'|+|\beta|=v$, and, for instance, $\beta_p\neq 0$ for some $0\le p\le l$. We claim that the polynomial

$$\tilde{P}(x,\zeta)=P^{(l_0)}(x,\zeta)+\frac{1}{2C_K|\xi'|}P_{(l_p)}(x,\zeta),$$

where $l_i=(0,\dots,0,1,0,\dots,0)$ and the constant C_K is from (1.3), is hyperbolic. For: Fix (x,ξ') and suppose the polynomial $P(x,\zeta_0,\xi')=P(\zeta_0)$ has the roots $\lambda_1<\lambda_2<\cdots<\lambda_q$ with multiplicities $r_1,\dots,r_q$, respectively. Then $\tilde{P}(\zeta_0)$ has roots at the points λ_i with multiplicities r_i-1. Moreover, for sufficiently small $\epsilon>0$ we have $\tilde{P}(\lambda_i+\epsilon)\tilde{P}(\lambda_{i+1}-\epsilon)<0$, since $P^{(l_0)}(\lambda_i+\epsilon)P^{(l_0)}(\lambda_{i+1}-\epsilon)<0$ and $\lim_{\epsilon\to 0}\frac{\tilde{P}(\lambda_i+\epsilon)}{P^{(l_0)}(\lambda_i+\epsilon)}=1+(2C_K|\xi'|)^{-1}\frac{\partial}{\partial x_p}\lambda_i(x,\xi')>0$. Consequently, $\tilde{P}(\zeta_0)$ has a root in the interval $(\lambda_i,\lambda_{i+1})$. This proves the claim.

Let us apply the induction hypothesis to $\tilde{\tilde{P}}=A\tilde{P}$, where

$$A=\left(m+\left((2C_K|\xi'|)^{-1}\frac{\partial}{\partial x_p}a_1(x,\xi')\right)\right)^{-1}.$$

Then we obtain

$$\begin{aligned}\left|\tilde{\tilde{P}}^{(\alpha+\gamma_0 l_0)}_{(\beta-l_p)}\right| &= \Bigg|\sum_{\sigma,\tau}\binom{\alpha+\gamma_0 l_0}{\sigma}\binom{\beta-l_p}{\tau}A^{(\sigma)}_{(\tau)}P^{(\alpha-\sigma+(\gamma_0+1)l_0)}_{(\beta-\tau-l_p)}\\ &\quad+\binom{\alpha+\gamma_0 l_0}{\sigma}\binom{\beta-l_p}{\tau}\left(\frac{A}{2C_K|\xi'|}\right)^{(\sigma)}_{(\tau)}P^{(\alpha-\sigma+\gamma_0 l_0)}_{(\beta-\tau)}\Bigg|\\ &\le C'_{K,k}\left|\tilde{\tilde{P}}\right|\frac{(1+|\xi'|)^{|\beta-l_p|-\frac{(|\alpha|+|\beta-l_p|-1)}{s}}}{(H(1+|\xi'|)^{\frac{1}{s}}+|\zeta|^{\frac{1}{s}})^{\gamma_0+1}}\\ &\le C_{K,k}|P|\frac{(1+|\xi'|)^{|\beta|-\frac{(|\alpha|+|\beta|-1)}{s}-1}}{(H(1+|\xi'|)^{\frac{1}{s}}+|\zeta|^{\frac{1}{s}})^{\gamma_0+1}}.\end{aligned}$$

All terms on the left-hand side of that inequality, apart from one, namely $\sigma=\tau=0$, satisfy the estimate in the induction hypothesis. Hence the required estimate holds also for this term. This completes the proof. □

1.2. Smooth structure on orbit spaces

One motivation for the second part of this book comes from the study of smooth structures on orbit spaces. Let us introduce the following notion of smoothness.

1.2.1. Definition. A *Frölicher space* is a triple $(X,\mathcal{C}_X,\mathcal{F}_X)$ consisting of a set X, a subset $\mathcal{C}_X$ of the set of all mappings $\mathbb{R}\to X$, and a subset $\mathcal{F}_X$ of the set of all functions $X\to\mathbb{R}$ such that:

(1) A function $f:X\to\mathbb{R}$ belongs to $\mathcal{F}_X$ if and only if $f\circ c\in C^\infty(\mathbb{R},\mathbb{R})$ for all $c\in\mathcal{C}_X$.
(2) A curve $c:\mathbb{R}\to X$ belongs to $\mathcal{C}_X$ if and only if $f\circ c\in C^\infty(\mathbb{R},\mathbb{R})$ for all $f\in\mathcal{F}_X$.

A mapping $\phi:X\to Y$ between two Frölicher spaces is called *smooth* if the following three equivalent conditions hold:

(3) For each $c\in\mathcal{C}_X$ the composition $\phi\circ c$ is in $\mathcal{C}_Y$.
(4) For each $f\in\mathcal{F}_Y$ the composition $f\circ\phi$ is in $\mathcal{F}_X$.
(5) For each $c\in\mathcal{C}_X$ and for each $f\in\mathcal{F}_Y$ the composition $f\circ\phi\circ c$ is in $C^\infty(\mathbb{R},\mathbb{R})$.

The set of all smooth functions from X to Y is denoted by $C^\infty(X,Y)$. Then we have $C^\infty(\mathbb{R},X) = \mathcal{C}_X$ and $C^\infty(X,\mathbb{R}) = \mathcal{F}_X$. Frölicher spaces and smooth mappings form a category.

If A is a subset of a Frölicher space X, with inclusion map $\iota_A : A \to X$, then A acquires a Frölicher structure by letting $\mathcal{C}_A$ be the set of all $c : \mathbb{R} \to A$ such that $\iota_A \circ c \in \mathcal{C}_X$.

1.2.2. The concept of Frölicher spaces was introduced, as *smooth spaces*, and studied by A. Frölicher in [**Frö80**] and [**Frö81**]. See also [**FK88**] and [**KM97**].

Note that a smooth manifold, with the usual notion of smooth curve into it and smooth real function on it, is a Frölicher space. This is a consequence of Boman's theorem [**Bom67**] which roughly speaking states that the smooth mappings on open subsets of $\mathbb{R}^n$ are exactly those mappings that map smooth curves to smooth curves.

1.2.3. Orbit spaces as Frölicher spaces. Let $\rho : G \to \mathrm{O}(V)$ be an orthogonal representation of a compact Lie group on a real finite dimensional Euclidean vector space. Then the algebra $\mathbb{R}[V]^G$ of G-invariant polynomials on V is finitely generated. Let $\sigma_1, \dots, \sigma_n$ be a system of homogeneous generators of $\mathbb{R}[V]^G$, and define $\sigma = (\sigma_1, \dots, \sigma_n) : V \to \mathbb{R}^n$. The image $\sigma(V)$ is a semialgebraic subset of $\mathbb{R}^n$. It turns out that σ induces a homeomorphism between the orbit space V/G and $\sigma(V)$. Hence we may identify V/G with $\sigma(V)$ and the orbit projection $V \to V/G$ with σ. For more details see section 8.2.

The orbit space $V/G = \sigma(V) \subseteq \mathbb{R}^n$ can be given the Frölicher subspace structure by setting

$$\mathcal{C}_{V/G} = \{c : \mathbb{R} \to \sigma(V) : \iota_{\sigma(V)} \circ c \text{ is smooth}\},$$

where $\iota_{\sigma(V)} : \sigma(V) \to \mathbb{R}^n$ is the inclusion map. It follows that the orbit map $\sigma : V \to \sigma(V)$ is smooth, i.e., $\sigma_* \mathcal{C}_V = \{\sigma \circ c : c \in \mathcal{C}_V\} \subseteq \mathcal{C}_{V/G}$. The question whether $\sigma_* \mathcal{C}_V = \mathcal{C}_{V/G}$ is exactly the problem tackled in the second part of this book: Given $c \in \mathcal{C}_{V/G}$, i.e., a smooth curve in $V/G = \sigma(V)$ (smooth as curve in $\mathbb{R}^n$), does there exist a $\bar{c} \in \mathcal{C}_V$ such that $\sigma \circ \bar{c} = c$, i.e., a smooth lift to the representation space V?

We shall see that this is not possible in general. That means the structure curves of a quotient of a Frölicher space need not be liftable as structure curves.

There is a second functional structure on V/G, see section 8.2. Namely, the algebra $C^\infty(V)^G$ of invariant smooth functions on V may be considered as algebra of smooth functions on V/G and, by Schwarz's theorem 8.2.5, identified with $\{f|_{\sigma(V)} : f \in C^\infty(\mathbb{R}^n)\}$. Thus, curves $c : \mathbb{R} \to \sigma(V) = V/G$ such that $f \circ c : \mathbb{R} \to \mathbb{R}$ is smooth for all $f \in C^\infty(V)^G$ correspond exactly to the smooth curves $c : \mathbb{R} \to \mathbb{R}^n$ with image in $\sigma(V)$, i.e., $c \in \mathcal{C}_{V/G}$. Now, suppose that $f \circ c$ is smooth for all $c \in \sigma_* \mathcal{C}_V$. Then $(f \circ \sigma) \circ \bar{c}$ is smooth for all $\bar{c} \in \mathcal{C}_V$, whence $f \circ \sigma$ is smooth, by Boman's theorem [**Bom67**]. Consequently, *the smooth structure as a Frölicher space of the orbit space V/G is unique.*

Part 1

Choosing roots of polynomials smoothly

CHAPTER 2

Continuity of the roots

The goal of this chapter is to establish a few results on the continuity of the roots of polynomials depending on their coefficients. All polynomials in this chapter are supposed to be over $\mathbb{C}$.

2.1. A first continuity theorem

2.1.1. PROPOSITION. *Let $P(x) = a_0 + a_1x + \cdots + a_nx^n$ be a monic, i.e., $a_n = 1$, polynomial over $\mathbb{C}$. Then, for each root w of P and for each $\epsilon > 0$ there is a $\delta > 0$ such that all monic polynomials $Q(x) = \sum_{i=0}^n b_ix^i$ with $|a_i - b_i| < \delta$, for $0 \le i \le n-1$, have a root z satisfying $|w - z| < \epsilon$.*

PROOF. Let $Q(x) = \sum_{i=0}^n b_ix^i = \prod_{i=1}^n (x - z_i)$ be a monic polynomial with roots $z_1, \dots, z_n$. For a root w of P we have

$$\prod_{i=1}^n (w - z_i) = Q(w) = Q(w) - P(w) = \sum_{i=0}^{n-1} (b_i - a_i)w^i,$$

whence

$$\min_{1\le i\le n} |w - z_i| \le \left(\sum_{i=0}^{n-1} |b_i - a_i||w|^i \right)^{\frac{1}{n}}.$$

The statement of the lemma is now obvious. □

2.1.2. The multiplicity of a root was no object in this first consideration. However, later on we will need the fact that, if w is a m-fold root of P and the coefficients of Q only differ slightly from those of P, then Q has m roots near w. Before we can prove this, we have to consider the following result, concerning moduli of roots, for preparation:

LEMMA. *Let $P(x) = a_0 + a_1x + \cdots + a_nx^n$ be a monic polynomial of degree n, and let $m \in \mathbb{N}$ with $m \le n$. Then P has at least m roots of modulus not exceeding*

$$2 \max_{0\le j\le m-1} |a_j|^{\frac{1}{n-j}}.$$

PROOF. We first prove a weaker statement. Let P belong to the following class of monic polynomials

$$\mathcal{M}_{m,n} = \left\{ \sum_{i=0}^n a_ix^i : a_n = 1, |a_j| \le 1 \text{ for } j = 0, \dots, m-1 \right\}.$$

Arranging the roots z_i of P such that $|z_1| \le \cdots \le |z_n|$, we assert that $|z_m| \le 2$.

If $|z_{m+1}| \le 2$, then the assertion is trivial. So let us suppose that $|z_{m+1}| > 2$. We want to factor out of P the roots $z_{m+1}, \dots, z_n$. Let z be one of them, and define

$$Q(x) = \sum_{i=0}^{n-1} b_ix^i = \frac{P(x)}{x - z}.$$

Then we find, by equating coefficients and putting $b_{-1} = b_n = 0$, that

$$a_i = -zb_i + b_{i-1} \quad \text{for } 0 \le i \le n.$$

Note that this corresponds exactly to Horner's algorithm. By solving this recurrence formula, we conclude that $b_j = -\sum_{i=0}^{j} \frac{a_i}{z^{j-i+1}}$ $(0 \le j \le n-1)$, whence, under our assumptions,

$$|b_j| \le \sum_{i=0}^{j} \frac{|a_i|}{|z|^{j-i+1}} \le \sum_{i=0}^{j} \frac{|a_i|}{2^{j-i+1}} \le \sum_{k=1}^{j+1} 2^{-k} < 1 \quad (0 \le j \le m-1).$$

That means that $Q \in \mathcal{M}_{m,n-1}$. Repeating this process, we see that

$$R(x) = \sum_{i=0}^{m} c_i x^i = \frac{P(x)}{(x - z_{m+1}) \cdots (x - z_n)}$$

belongs to $\mathcal{M}_{m,m}$. Next, let w be a root of R, so $w^m = -\sum_{i=0}^{m-1} c_i w^i$. Thus, in the case where $|w| > 1$, we have

$$|w| \le \sum_{i=0}^{m-1} |c_i||w|^{i-m+1} \le \sum_{k=0}^{\infty} |w|^{-k} = \frac{1}{1 - |w|^{-1}}$$

which implies that $|w| \le 2$. Hence, we have shown that, in any case, $|z_m| \le 2$. The above assertion is verified.

Now, let us deduce the statement of the lemma. Set

$$\lambda = \max_{0 \le j \le m-1} |a_j|^{\frac{1}{n-j}}.$$

If $\lambda = 0$, then $a_0 = a_1 = \cdots = a_{m-1} = 0$, so 0 is an m-fold root of P, and the assertion of the lemma is trivially satisfied. Suppose $\lambda > 0$. Then

$$\lambda^{-n} P(\lambda x) = \sum_{i=0}^{n} \lambda^{i-n} a_i x^i$$

belongs to $\mathcal{M}_{m,n}$, since, for $0 \le i \le m-1$, we have $|a_i| \le \max_{0 \le j \le m-1} |a_j|^{\frac{n-i}{n-j}}$ which implies

$$\lambda^{i-n}|a_i| = \left(\max_{0 \le j \le m-1} |a_j|^{\frac{1}{n-j}} \right)^{i-n} |a_i| = \left(\max_{0 \le j \le m-1} |a_j|^{\frac{n-i}{n-j}} \right)^{-1} |a_i| \le 1.$$

By the above, $\lambda^{-n} P(\lambda x)$ has at least m roots of modulus not exceeding 2, and, hence, P has at least m roots of modulus not exceeding 2λ. $\square$

Now we are prepared to show the following deeper theorem on the continuity of roots of polynomials as functions of the coefficients. This theorem is basically due to Ostrowski [**Ost40**]. We follow the presentation in [**RS02**].

THEOREM. *Let*

$$P(x) = \sum_{i=0}^{n} a_i x^i = \prod_{j=1}^{p} (x - x_j)^{m_j} \qquad (m_1 + \cdots + m_p = n)$$

be a monic polynomial of degree n with distinct roots $x_1, \ldots, x_p$ of multiplicities $m_1, \ldots, m_p$. Then, given a positive $\epsilon < \min_{1 \le i < j \le p} \frac{|x_i - x_j|}{2}$, there exists a positive δ so that any monic polynomial $Q(x) = \sum_{i=0}^{n} b_i x^i$ whose coefficients satisfy $|a_i - b_i| < \delta$, for $0 \le i \le n-1$, has exactly m_j roots in the disk

$$D(x_j; \epsilon) = \{z \in \mathbb{C} : |z - x_j| \le \epsilon\} \quad (1 \le j \le p).$$

More precisely: Let

$$A = \max\left\{1, 2|a_i|^{\frac{1}{n-i}} : 0 \le i \le n-1\right\},$$

and let the roots of P be denoted by $z_1, \ldots, z_n$ where an m-fold root is now listed m times. Then, for sufficiently small $\delta > 0$, there exists a numbering of the roots of Q as $w_1, \ldots, w_n$ such that

$$\max_{1 \le i \le n} |w_i - z_i| \le 4A\delta^{\frac{1}{n}}.$$

PROOF. Expansions (via Taylor's formula) of the polynomials P and Q at x_j yield

$$P(x + x_j) = \sum_{i=0}^{n} a_{j,i} x^i \quad \text{and} \quad Q(x + x_j) = \sum_{i=0}^{n} b_{j,i} x^i$$

where

$$a_{j,i} = \frac{1}{i!} P^{(i)}(x_j) = \frac{1}{i!} \sum_{k=i}^{n} \frac{k!}{(k-i)!} a_k x_j^{k-i} = \sum_{k=i}^{n} \binom{k}{i} a_k x_j^{k-i}$$

and as well

$$b_{j,i} = \sum_{k=i}^{n} \binom{k}{i} b_k x_j^{k-i}.$$

Note that $a_{j,n} = a_n = b_n = b_{j,n} = 1$. Furthermore, since x_j is an m_j-fold root of P, we have $a_{j,0} = \cdots = a_{j,m_j-1} = 0$, and, therefore,

$$b_{j,l} = b_{j,l} - a_{j,l} = \sum_{k=l}^{n-1} \binom{k}{l} (b_k - a_k) x_j^{k-l} \quad (0 \le l \le m_j - 1).$$

Now, applying the forgoing lemma to $Q(x + x_j)$ (viewed as polynomial in x) with $m = m_j$ and introducing

$$\rho_j = 2 \max_{0 \le l \le m_j - 1} \left(\sum_{k=l}^{n-1} \binom{k}{l} |b_k - a_k| |x_j|^{k-l} \right)^{\frac{1}{n-l}},$$

we find that $D(x_j; \rho_j)$ contains at least m_j roots of Q. By choosing δ sufficiently small, the radii ρ_j can all be made smaller than $\epsilon < \min_{1 \le i < j \le p} \frac{|x_i - x_j|}{2}$. Then the disks $D(x_1; \rho_1), \ldots, D(x_p; \rho_p)$ are disjoint. Thus, each $D(x_j; \rho_j)$ must contain exactly m_j roots.

To verify the supplement in the theorem, it suffices to show that $4A\delta^{\frac{1}{n}}$ is an upper bound for the radii ρ_j, at least for small $\delta > 0$. By the above lemma, the moduli of the roots of P are bounded by A. Since $A \ge 1$ and $\binom{k}{l} < 2^k$, we find that

$$\sum_{k=l}^{n-1} \binom{k}{l} |b_k - a_k| |x_j|^{k-l} < 2^n \delta A^{n-l}.$$

Hence, for $0 < \delta < 2^{-n}$, we have

$$\rho_j < 2 \max_{0 \le l \le m_j - 1} (2^n \delta)^{\frac{1}{n-l}} A \le 4A\delta^{\frac{1}{n}}$$

which concludes the proof. □

REMARK. In view of the second statement in the previous theorem, we may say that the roots of a polynomial of degree n, as functions of the coefficients, satisfy a local Hölder condition of order $\frac{1}{n}$.

2.2. Rouché's theorem and an application

Another possibility to get results on the continuity of roots of polynomials is the application of Rouché's theorem. We shall first derive Rouché's theorem. To start with let us recall a few results from complex analysis.

2.2.1. Let $a \in \mathbb{C}$, and let $\gamma : [0,1] \to \mathbb{C}\backslash\{a\}$ be a *circuit*, i.e., a closed path such that γ is a primitive of a *regulated function*, i.e., a function having one-sided limits at every point. We define the *index of a with respect to γ* as

$$j(a;\gamma) := \frac{1}{2\pi i}\int_\gamma \frac{dz}{z-a}.$$

This number is always an integer: Consider the function $h : [0,1] \to \mathbb{C}$ with $h(t) = \int_0^t \frac{\gamma'(s)}{\gamma(s)-a}ds$. It has a derivative equal to $h'(t) = \frac{\gamma'(t)}{\gamma(t)-a}$ except at the points of an at most denumerable subset of $[0,1]$. Hence, if we put $g(t) := e^{-h(t)}(\gamma(t)-a)$, we find that $g'(t) = 0$, except in an at most denumerable subset of $[0,1]$. It follows that g is constant, and this constant is easily found to be $\gamma(0) - a$, thus, $e^{h(t)} = \frac{\gamma(t)-a}{\gamma(0)-a}$. But we have $\gamma(0) = \gamma(1)$, whence $e^{h(1)} = 1$. So $h(1) = 2n\pi i$ for an integer n. And we have $n = \frac{1}{2\pi i}h(1) = \frac{1}{2\pi i}\int_0^1 \frac{\gamma'(s)}{\gamma(s)-a}ds = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z-a} = j(a;\gamma)$.

From Cauchy's theorem (well-known from complex analysis, see e.g. [**Die60**]) it follows that, if γ_1 and γ_2 are circuits in $\mathbb{C}\backslash\{a\}$ which are homotopic in that set, then they have the same index with respect to a.

Consider in particular the circuit $\gamma_0(t) = a + re^{2\pi it}$. Then

$$j(a;\gamma_0) = \frac{1}{2\pi i}\int_{\gamma_0}\frac{dz}{z-a} = \frac{1}{2\pi i}\int_0^1 \frac{2\pi i r e^{2\pi it}}{re^{2\pi it}}dt = 1.$$

Since any simple circuit γ which contains a is homotopic in the set $\mathbb{C}\backslash\{a\}$ to γ_0, it follows that the index of a with respect to a simple circuit γ containing a equals 1.

Claim (1). *The index $j(x;\gamma)$ is constant in each connected component of the complement A of $\gamma([0,1])$.*

Proof. The set $\gamma([0,1])$ is compact, and so A is open. We will show first that the mapping $x \mapsto j(x;\gamma)$ is continuous in the open set A. For: Observe that the index of $x+h$ with respect to γ (when $x+h \notin \gamma([0,1])$) is equal to the index of x with respect to the circuit $\gamma_1 : t \mapsto \gamma(t) - h$:

$$\int_\gamma \frac{dz}{z-(x+h)} = \int_0^1 \frac{\gamma'(t)}{(\gamma(t)-h)-x}dt = \int_{\gamma_1}\frac{dz}{z-x}.$$

But, if B is a ball centered at x with radius r which is contained in A, then $\phi : [0,1]\times[0,1] \to \mathbb{C}$ with $\phi(t,s) = \gamma(t) - sh$ is a loop homotopy of γ to γ_1 in $\mathbb{C}\backslash\{x\}$, as long as $|h| < r$. Therefore, by the remark above, we have $j(x+h;\gamma) = j(x;\gamma)$, and hence continuity.

Now, since $x \mapsto j(x;\gamma)$ is continuous and the target space $\mathbb{Z}$ is discrete, the claim follows. □

Claim (2). *If a circuit γ is contained in a closed ball $D = \{z \in \mathbb{C} : |z-a| \le r\}$, then $j(z;\gamma) = 0$ for any point z exterior to D.*

Proof. Suppose the circuit $\gamma : [0,1] \to \mathbb{C}$ is contained in D, and assume that $|\gamma'(t)| \le M$ for all $t \in [0,1]$ (which is possible, since γ' is a regulated function). Let z be exterior to D, i.e., $|z-a| > r$. By definition, we have

$$j(z;\gamma) = \frac{1}{2\pi i}\int_\gamma \frac{dx}{x-z} = \frac{1}{2\pi i}\int_0^1 \frac{\gamma'(t)}{\gamma(t)-z}dt.$$

As $|\gamma(t)-a| \le r$ for all $t \in [0,1]$, we find $|z-a| \le |z-\gamma(t)|+|\gamma(t)-a| \le |\gamma(t)-z|+r$ for any $t \in [0,1]$. Consequently,

$$|j(z;\gamma)| = \frac{1}{2\pi}\left|\int_0^1 \frac{\gamma'(t)}{\gamma(t)-z}dt\right| \le \frac{1}{2\pi}\int_0^1 \frac{|\gamma'(t)|}{|\gamma(t)-z|}dt \le \frac{1}{2\pi}\frac{M}{|z-a|-r}.$$

If $|z-a|$ is large enough, then the right-hand side of the last inequality is less than 1, and, since $j(z;\gamma)$ is an integer, we conclude that $j(z;\gamma)=0$. The statement follows by claim 1, since the exterior of D is connected. $\square$

Let A be an open subset of $\mathbb{C}$, a an isolated point of $\mathbb{C}\backslash A$, and $r>0$ such that all points of the ball $B=\{z\in\mathbb{C}:|z-a|\le r\}$ except a belong to A. If $f:A\to\mathbb{C}$ is holomorphic, then for $0<|z-a|<r$, we have the Laurent series (see e.g. [**Die60**])

$$f(z)=\sum_{n=0}^{\infty}c_n(z-a)^n+\sum_{n=1}^{\infty}d_n(z-a)^{-n},$$

where both series are convergent for $0<|z-a|<r$, and

$$c_n=\frac{1}{2\pi i}\int_\gamma\frac{f(x)}{(x-a)^{n+1}}dx,\qquad d_n=\frac{1}{2\pi i}\int_\gamma(x-a)^{n-1}f(x)dx,$$

where γ is the circuit $t\mapsto a+re^{it}$, for $0\le t\le 2\pi$.

Consider the series $u(x)=\sum_{n=1}^{\infty}d_nx^n$ and call $u((z-a)^{-1})$ the *singular part of f at a*. When $u\ne 0$, we say that a is an *isolated singular point of f*. If u is a polynomial of degree $n\ge 1$, then a is called a *pole of order n of f*; otherwise, i.e., if infinitely many coefficients in u do not vanish, then a is called an *essential singularity of f*.

We define the *order of f at a* denoted by $\omega(a;f)$ as follows: Put $\omega(a;f)=-\infty$, if a is an essential singularity of f; $\omega(a;f)=-n$, if a is a pole of order n; $\omega(a;f)=m$, if $f\ne 0$, $u=0$, and in the power series $\sum_{n=0}^{\infty}c_n(z-a)^n$, equal to $f(z)$ for $0<|z-a|<r$, m is the smallest integer for which $c_m\ne 0$; finally, we set $\omega(a;0)=+\infty$. For $\omega(a;f)=m>0$, we also say that a is a *zero of order m of f*.

Finally, note that the coefficient $d_1=\frac{1}{2\pi i}\int_\gamma f(x)dx$ in the above presentation is called the *residue of f at a*.

Let A be an open subset of $\mathbb{C}$ and S a subset of A, all points of which are isolated. A mapping $f:A\backslash S\to\mathbb{C}$ is called *meromorphic in A*, if it is holomorphic in $A\backslash S$ and no point of S is an essential singularity. If f is meromorphic in A, S is the set of its poles, T the set of its zeros, then all the points in $S\cup T$ are isolated. Namely, if $a\in A$ and $\omega(a;f)=h$, then $f(z)=(z-a)^hf_1(z)$ for $0<|z-a|<r$ with r sufficiently small, where f_1 is holomorphic for $|z-a|<r$ and $f_1(a)\ne 0$. Moreover, we can find a number $0<r'<r$ such that $f_1(z)\ne 0$ for $0<|z-a|<r'$. So we have found a neighborhood U of a such that $U\backslash\{a\}$ contains no point of $S\cup T$.

In particular, $\frac{1}{f}$ is meromorphic in A as well, and S is its set of zeros and T is its set of poles. With the same notion as above we have $f'(z)=h(z-a)^{h-1}f_1(z)+(z-a)^hf_1'(z)$ for $0<|z-a|<r$, thus,

$$\frac{f'(z)}{f(z)}=\frac{h}{z-a}+\frac{f_1'(z)}{f_1(z)}$$

is holomorphic for $0<|z-a|<r'$ and has order 0 at a if $h=0$, order -1 and residue h at a if $h\ne 0$.

Proposition. *Let $A\subseteq\mathbb{C}$ be a simply connected domain, and suppose $f:A\to\mathbb{C}$ is a meromorphic function on A. Let S be the set of poles of f and T the set of zeros of f. Consider a circuit $\gamma:[0,1]\to A\backslash(S\cup T)$. Then, for the circuit $\Gamma=f\circ\gamma$ we have*

$$j(0;\Gamma)=\sum_{a\in S\cup T}j(a;\gamma)\omega(a;f).$$

PROOF. We have, by definition,

$$j(0;\Gamma) = \frac{1}{2\pi i}\int_\Gamma \frac{dx}{x} = \frac{1}{2\pi i}\int_0^1 \frac{\Gamma'(t)}{\Gamma(t)}dt = \frac{1}{2\pi i}\int_0^1 \frac{f'(\gamma(t))\gamma'(t)}{f(\gamma(t))}dt = \frac{1}{2\pi i}\int_\gamma \frac{f'(z)}{f(z)}dz.$$

By the theorem of residues (see e.g. [**Die60**]) and the discussion before the proposition, it follows that

$$\int_\gamma \frac{f'(z)}{f(z)}dz = 2\pi i \sum_{a\in S\cup T} j(a;\gamma)\omega(a;f).$$

So the proof is complete. □

2.2.2. THEOREM (Rouché's theorem). *Let $A \subseteq \mathbb{C}$ be a simply connected domain, and suppose $f, g : A \to \mathbb{C}$ are holomorphic functions on A. Let T be the (at most denumerable) set of zeros of f in A, T' the set of zeros of $f+g$ in A, and let $\gamma : [0,1] \to A\backslash T$ be a circuit. Then, if $|f(z)| > |g(z)|$ for $z \in \gamma([0,1])$, the function $f+g$ has no zeros on $\gamma([0,1])$, and we have*

$$\sum_{a\in T} j(a;\gamma)\omega(a;f) = \sum_{b\in T'} j(b;\gamma)\omega(b;f+g). \tag{2.1}$$

In particular, if γ is a simple circuit, then the function $f+g$ has as many roots (with multiplicities) in the interior of $\gamma([0,1])$ as f.

PROOF. Obviously, the function $f+g$ does not vanish on $\gamma([0,1])$: If there were a $z \in \gamma([0,1])$ such that $f(z)+g(z) = 0$, then $|f(z)| = |g(z)|$, in contradiction to our assumption. The function $h = \frac{f+g}{f}$ is defined on $A\backslash T$ and is meromorphic on A. We have

$$h' = \frac{(f'+g')f - (f+g)f'}{f^2} = \frac{fg' - f'g}{f^2}$$

on $A\backslash T$, and, therefore,

$$\frac{f'}{f} + \frac{h'}{h} = \frac{f'(f+g)}{f(f+g)} + \frac{fg'-f'g}{f(f+g)} = \frac{f'f+fg'}{f(f+g)} = \frac{(f+g)'}{f+g}$$

on $A\backslash(T\cup T')$. By our initial observation, we find $\gamma([0,1]) \subseteq A\backslash(T\cup T')$, and thus

$$\int_\gamma \frac{(f(z)+g(z))'}{f(z)+g(z)}dz = \int_\gamma \frac{f'(z)}{f(z)}dz + \int_\gamma \frac{h'(z)}{h(z)}dz. \tag{2.2}$$

By proposition 2.2.1, it suffices to prove that the index of 0 with respect to the circuit $\Gamma = h\circ\gamma$ is 0, since then the right-most term in (2.2) vanishes and the equation (2.2) turns into (2.1). Since $r := \sup_{z\in\gamma([0,1])}\left|\frac{g(z)}{f(z)}\right| < 1$ by assumption, we have $|h(z)-1| = \left|\frac{g(z)}{f(z)}\right| < r < 1$ for all $z \in \gamma([0,1])$. So the circuit Γ lies in the ball $\{z \in \mathbb{C} : |z-1| < r\}$, and as 0 is exterior to this ball, the result follows from claim 2.

If γ is simple, then the indices $j(a;\gamma)$ (resp. $j(b;\gamma)$) in (2.1) are either 0 if a (resp. b) is exterior to γ, or 1 if a (resp. b) is interior to γ. Hence, also the additional assertion is proved. □

2.2.3. Continuity of the roots of an equation depending on parameters.

THEOREM. *Let A be an open set in $\mathbb{C}$, $f : \mathbb{R}\times A \to \mathbb{C}$ a continuous function, such that for each $t \in \mathbb{R}$, $z \mapsto f(t,z)$ is holomorphic on A and does not vanish identically. If the equation $f(t_0,z) = 0$ has a root $z_0 \in A$ of multiplicity r, then, in a sufficiently small neighborhood of $(t_0,z_0) \in \mathbb{R}\times A$, the equation $f(t,z) = 0$ has r (with multiplicities) roots $z_j = z_j(t)$, $1 \le j \le r$, and $\lim_{t\to t_0} z_j(t) = z_0$, for $1 \le j \le r$.*

Proof. By assumption, the map $z \mapsto f(t_0, z)$ is holomorphic on A, it does not vanish identically, and $z_0 \in A$. Therefore, z_0 is an isolated root of $f(t_0, z) = 0$, and we may choose a small circle $\gamma : [0, 1] \to \mathbb{C}$ in A with center z_0, such that z_0 is the only root of $f(t_0, z) = 0$ lying in the interior of $\gamma([0, 1])$, and no root is lying on $\gamma([0, 1])$.

Let $m = \min_{z\in\gamma([0,1])} |f(t_0, z)|$, then $m > 0$, since $\gamma([0, 1])$ is compact and $z \mapsto f(t_0, z)$ is continuous. By continuity of f in both variables, for each $z \in \gamma([0, 1])$ there is a neighborhood U_z of z contained in A and a neighborhood V_z of t_0 in $\mathbb{R}$ such that $|f(t, w) - f(t_0, z)| \le |f(t, w) - f(t, z)| + |f(t, z) - f(t_0, z)| < \frac{m}{2}$ for all $w \in U_z$ and $t \in V_z$. The compact set $\gamma([0, 1])$ can be covered by finitely many U_{z_k}. Then $V = \bigcap_k V_{z_k}$ defines a neighborhood of t_0 in $\mathbb{R}$ such that for all $z \in \gamma([0, 1])$ and all $t \in V$

$$|f(t, z) - f(t_0, z)| < \min_{z\in\gamma([0,1])} |f(t_0, z)| \le |f(t_0, z)|.$$

We can apply Rouché's theorem 2.2.2. Consequently, for $t \in V$ the equation $f(t, z) = 0$ has as many roots in the interior of $\gamma([0, 1])$ as $f(t_0, z) = 0$ has. So there are r (with multiplicities) roots $z_j = z_j(t)$ of $f(t, z) = 0$ in a neighborhood of (t_0, z_0), and $\lim_{t\to t_0} z_j(t) = z_0$ for all j, since we may shrink the circle $\gamma([0, 1])$ to the point z_0. $\square$

Remark. The parameter space $\mathbb{R}$ in the theorem can be replaced by any metric space.

2.3. Continuous parameterizations of the roots

The continuity formulated in theorem 2.1.2 and theorem 2.2.3 is the continuity of the roots as a whole. It is a different question whether it is possible to define n single-valued continuous functions which parameterize the roots of a family of polynomials of degree n.

Such a parameterization is impossible in general:

Example. Consider $P(\alpha)(z) = z^2 - \alpha$ for $\alpha \in \mathbb{C}$. It is not possible to find two single-valued continuous functions representing the two roots $\pm\sqrt{\alpha}$ in a domain of the complex plane containing the branch point 0.

A continuous parameterization of the single roots is possible if either the parameter space is $\mathbb{R}$ or the roots are always real. This will be shown in the following two theorems. See [**AKLM98**] and [**Kat76**].

2.3.1. Theorem. *Consider a curve of monic polynomials*

$$P(t)(z) = z^n - a_1(t)z^{n-1} + \cdots + (-1)^n a_n(t),$$

with continuous coefficients $a_i : \mathbb{R} \to \mathbb{C}$, $1 \le i \le n$. Then there exist n continuous functions $z_i : \mathbb{R} \to \mathbb{C}$, $1 \le i \le n$, which parameterize the roots of P.

Proof. The repeated roots of a polynomial of degree n form an unordered n-tuple of complex numbers. Two such n-tuples $z = (z_1, \dots, z_n)$ and $z' = (z'_1, \dots, z'_n)$ may be considered close to each other if, for suitable numbering of their elements, the $|z_i - z'_i|$ are small for all $1 \le i \le n$. We can define the distance between two such n-tuples by

$$\operatorname{dist}(z, z') = \min \max_{1\le i\le n} |z_i - z'_i|, \tag{2.3}$$

where the minimum is taken over all possible re-numberings of the elements of one of the n-tuples. In this language, theorem 2.1.2 and theorem 2.2.3 imply that the unordered n-tuple $z(t)$ consisting of the repeated roots of $P(t)$ changes with t continuously.

Now we prove, by induction on the degree n, that there exist n continuous functions $z_i(t)$, $1 \le i \le n$, which constitute the n-tuple $z(t)$ for each $t \in \mathbb{R}$. This implies the theorem. If $n = 1$ there is nothing to prove. So let us assume $n > 1$. Let E be the set of points $t \in \mathbb{R}$ such that all roots of $P(t)$ coincide and thus all elements of $z(t)$ are identical. Then E is closed and $\mathbb{R}\backslash E$ is the disjoint union of at most countably many open intervals I_k (each of which maximally connected). Fix some I_k and let $t_0 \in I_k$. Since not all elements of $z(t_0)$ are identical, they may be divided in two disjoint subsets with n_1 and n_2 elements ($n_1+n_2 = n$). In other words, $z(t_0)$ is composed of an n_1-tuple and an n_2-tuple with separate elements. The continuity of $z(t)$ guarantees that for t near t_0 the n-tuple $z(t)$ consists likewise of an n_1-tuple and an n_2-tuple each of which is continuous in t. By induction hypothesis, these tuples can be represented near t_0 by families of continuous functions which, taken together, represent $z(t)$.

So any $t \in I_k$ has a neighborhood such that the n-tuple z is constituted by n continuous functions z_i in this neighborhood. We want to show that this holds globally on I_k. Suppose the continuous functions z_i representing z are maximally defined on $(a, b) \subseteq I_k$ with $b \in I_k$. Then there exist n continuous functions z'_i which represent z near b and which are defined for some $t_1 < b$. Since $(z_1(t_1), \dots, z_n(t_1)) = (z'_1(t_1), \dots, z'_n(t_1))$ as unordered tuples, after a suitable renumbering of $(z'_1, \dots, z'_n)$, we have $z_i(t_1) = z'_i(t_1)$, $1 \le i \le n$. It follows that the functions $\hat{z}_i$, defined by $\hat{z}_i(t) := z_i(t)$ for $t \le t_1$ and $\hat{z}_i(t) := z'_i(t)$ for $t \ge t_1$, are continuous and represent z on a larger interval than (a, b).

We have shown that on each I_k there exists a continuous parameterization z_i^k, $1 \le i \le n$, of the elements of z. For $t \in E$, the n-tuple $z(t)$ consists of n identical elements $\zeta(t)$. Let us define n functions z_i, $1 \le i \le n$, on $\mathbb{R}$ by

$$z_i(t) := \begin{cases} z_i^k(t) & \text{if } t \in I_k \\ \zeta(t) & \text{if } t \in E. \end{cases}$$

These n functions represent z and are continuous on $\mathbb{R}$, which is now obvious. $\square$

Remark. If we endow the space $\mathcal{Z}_n$ of unordered n-tuples of complex numbers with the distance function defined in (2.3) on the one hand, and the space $\mathcal{P}_n = \{(a_1, \dots, a_n) \in \mathbb{C}^n : P(z) = z^n - a_1 z^{n-1} + \dots + (-1)^n a_n\} \cong \mathbb{C}^n$ of monic complex polynomials P of degree n with the metric $\max_{1 \le i \le n} |a_i - b_i|$ on the other hand, then the map $\mathcal{P}_n \to \mathcal{Z}_n$ which assigns to each polynomial $P \in \mathcal{P}_n$ the unordered n-tuple of its roots $z(P) \in \mathcal{Z}_n$ is a homeomorphism between the corresponding metric spaces. Cf. [**ĆM06**].

2.3.2. Theorem. *For a polynomial*

$$P(x) = x^n - a_1(P)x^{n-1} + \dots + (-1)^n a_n(P),$$

with all roots real, let $x_1(P) \le x_2(P) \le \dots \le x_n(P)$ *be the roots of* P*, increasingly ordered. Denote by* $\mathrm{Hyp}_n = \{(a_1, \dots, a_n) \in \mathbb{R}^n : P(x)$ *has all roots real*$\}$ *the space of all monic polynomials of degree* n *with all roots real. Then all roots* $x_i : \mathrm{Hyp}_n \to \mathbb{R}$ *are continuous.*

Proof. First we show that x_1 is continuous. Consider an arbitrary $P_0 \in \mathrm{Hyp}_n$. We have to show that for every $\epsilon > 0$ there exists some $\delta > 0$ such that for all $P \in \mathrm{Hyp}_n$ with $|P - P_0| < \delta$ there is a root $x(P)$ of P with $x(P) < x_1(P_0) + \epsilon$ and for all roots $x(P)$ of P we have $x(P) > x_1(P_0) - \epsilon$ (by the ordering of the roots). Without loss of generality we may assume that $x_1(P_0) = 0$.

We make induction on the degree n of P. For $n = 1$ the statement is evidently true. Let us assume that it holds whenever the degree is strictly smaller than n. By the splitting lemma 4.2.3, we can factorize $P = P_1(P) \cdot P_2(P)$, where $P_1(P_0)$ has all roots equal to $x_1(P_0) = 0$ and $P_2(P_0)$ has all roots greater than 0 and both polynomials have coefficients which depend real analytically on P. The degree of $P_2(P)$ is now smaller than n, consequently, by induction hypothesis, the roots of $P_2(P)$ are continuous and thus larger than $x_1(P_0) - \epsilon$ for P near P_0.

Since 0 was the smallest root of P_0, what remains to show is that for all $\epsilon > 0$ there exists a $\delta > 0$ such that for $|P - P_0| < \delta$ any root x of $P_1(P)$ satisfies $|x| < \epsilon$. Suppose there is a root x of $P_1(P)$ with $|x| \geq \epsilon$. Let n_1 denote the degree of P_1. From $P_1(P)(x) = 0$ we obtain

$$-x^{n_1} = \sum_{k=1}^{n_1} (-1)^k a_k(P_1(P)) x^{n_1 - k},$$

whence

$$\epsilon \leq |x| = \left| \sum_{k=1}^{n_1} (-1)^k a_k(P_1(P)) x^{1-k} \right| \leq \sum_{k=1}^{n_1} |a_k(P_1(P))| |x|^{1-k} < \sum_{k=1}^{n_1} \frac{\epsilon^k}{n_1} \epsilon^{1-k} = \epsilon,$$

provided that $n_1 |a_k(P_1(P))| < \epsilon^k$, which is true for $P_1(P)$ near P_0, since $a_k(P_1(P_0)) = 0$ for $1 \leq k \leq n_1$. This a contradiction and therefore x_1 is continuous.

To prove the continuity of the remaining roots $x_2(P) \leq \cdots \leq x_n(P)$ we use Horner's algorithm. We factorize $P(x) = (x - x_1(P)) \cdot P_3(P)(x)$, where $P_3(P)$ has the roots $x_2(P) \leq \cdots \leq x_n(P)$. Then there are the following relations between the coefficients $a_1, \ldots, a_n$ of P and those of $P_3(P)$, say $b_1, \ldots, b_{n-1}$:

$$a_n = b_{n-1} x_1,\ a_{n-1} = b_{n-1} + b_{n-2} x_1,\ \ldots,\ a_2 = b_2 + b_1 x_1,\ a_1 = b_1 + x_1.$$

It follows that the coefficients $b_1, \ldots, b_{n-1}$ of $P_3(P)$ are again continuous and so we can proceed by induction on the degree of P. Hence the theorem is proved. $\square$

CHAPTER 3

Hyperbolic polynomials

3.1. The space of hyperbolic polynomials

Let us introduce the following notion. A monic polynomial in one variable with real coefficients is called *hyperbolic*, if all its roots are real.

There is an elegant description of the space Hyp_n of hyperbolic polynomials with degree n as semialgebraic set in $\mathbb{R}^n$. We shall derive it in theorem 3.1.2.

3.1.1. The Bezoutiant. Let

$$P(x) = x^n - a_1 x^{n-1} + \cdots + (-1)^n a_n$$

be a monic polynomial with real coefficients $a_1, \ldots, a_n$ and roots $x_1, \ldots, x_n \in \mathbb{C}$. By *Vieta's formulas*, we know that $a_i = \sigma_i(x_1, \ldots, x_n)$, where $\sigma_1, \ldots, \sigma_n$ are the *elementary symmetric functions* in n variables:

$$\sigma_i(x_1, \ldots, x_n) = \sum_{1 \le j_1 < \cdots < j_i \le n} x_{j_1} \cdots x_{j_i}.$$

Denote by s_i ($i \in \mathbb{N}_0$) the *Newton polynomials* $\sum_{j=1}^n x_j^i$ which are related to the elementary symmetric functions by

$$s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 - \cdots + (-1)^{k-1} s_1 \sigma_{k-1} + (-1)^k k \sigma_k = 0 \quad (k \ge 1). \tag{3.1}$$

This relation is easily derived: The generating function for the s_i is

$$S(t) = \sum_{i \ge 1} s_i t^{i-1} = \sum_{i \ge 1} \sum_{j \ge 1} x_j^i t^{i-1} = \sum_{j \ge 1} \frac{x_j}{1 - x_j t} = \sum_{j \ge 1} \frac{d}{dt} \log \frac{1}{1 - x_j t}$$

such that

$$S(-t) = \frac{d}{dt} \log \Sigma(t) = \frac{\Sigma'(t)}{\Sigma(t)}, \tag{3.2}$$

where $\Sigma(t) = \sum_{i \ge 0} \sigma_i t^i = \prod_{j \ge 1} (1 + x_j t)$, ($\sigma_0 = 1$), is the generating function for the σ_i. Now (3.2) implies (3.1).

The relation (3.1) corresponds to a polynomial diffeomorphism ψ^n with $s^n = \psi^n \circ \sigma^n$, where we define $\sigma^n := (\sigma_1, \ldots, \sigma_n) : \mathbb{R}^n \to \mathbb{R}^n$ and $s^n := (s_1, \ldots, s_n) : \mathbb{R}^n \to \mathbb{R}^n$. Note that the Jacobian of s^n is $n!$-times the Vandermonde determinant:

$$\det(ds^n(x)) = n! \prod_{i > j} (x_i - x_j) = n! \mathrm{Van}(x).$$

Even the derivative $ds^n(x)$ itself equals the Vandermonde matrix up to factors i in the i-th row. Furthermore, we have

$$\det(d\psi^n(x)) = (-1)^{\frac{n(n+3)}{2}} n! = (-1)^{\frac{n(n-1)}{2}} n!,$$

and consequently

$$\det(d\sigma^n(x)) = \prod_{i < j} (x_i - x_j).$$

Let us consider the so-called *Bezoutiant*

$$B := \begin{pmatrix} s_0 & s_1 & \dots & s_{n-1} \\ s_1 & s_2 & \dots & s_n \\ \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_n & \dots & s_{2n-2} \end{pmatrix}.$$

Note that the term Bezoutiant is used for different objects in the mathematical literature, compare with [**RS02**]. Denote by B_k the minor formed by the first k rows and columns of B. From

$$B_k(x) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{k-1} & x_2^{k-1} & \dots & x_n^{k-1} \end{pmatrix} \cdot \begin{pmatrix} 1 & x_1 & \dots & x_1^{k-1} \\ 1 & x_2 & \dots & x_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_n^{k-1} \end{pmatrix}$$

it follows that

$$\Delta_k(x) := \det(B_k(x)) = \sum_{i_1<i_2<\dots<i_k} (x_{i_1}-x_{i_2})^2 \cdots (x_{i_1}-x_{i_k})^2 \cdots (x_{i_{k-1}}-x_{i_k})^2, \tag{3.3}$$

since for $(k\times n)$-matrices A one has $\det(AA^\top) = \sum_{i_1<\dots<i_k} \det(A_{i_1,\dots,i_k})^2$, where $A_{i_1,\dots,i_k}$ is the minor of A with indicated rows and columns, by lemma 3.1.3 below. Since the polynomials Δ_k are symmetric, we have $\Delta_k = \tilde{\Delta}_k \circ \sigma^n$ for unique polynomials $\tilde{\Delta}_k$. Similarly we find a unique symmetric $(n\times n)$-matrix $\tilde{B}$ with $B = \tilde{B}\circ\sigma^n$. Compare with 3.1.4.

3.1.2. The following theorem is Sylvester's version of a theorem of Sturm giving a nice characterization of the space Hyp_n of hyperbolic polynomials of degree n. The proof presented here is due to Procesi [**Pro78**].

THEOREM. *Let P be a monic polynomial of degree n with real coefficients $a_1,\dots,a_n$. Then the following statements are equivalent:*

(1) *P is hyperbolic.*
(2) *$\tilde{B}(P)$ is positive semidefinite.*
(3) *All determinants of principal (i.e. symmetric) minors of $\tilde{B}(P)$ are non-negative; in particular $\tilde{\Delta}_k(P) = \tilde{\Delta}_k(a_1,\dots,a_n) \ge 0$ for $1\le k\le n$.*

Moreover, the rank of $\tilde{B}(P)$ equals the number of distinct roots of P and its signature equals the number of distinct real roots.

PROOF. The equivalence of (2) and (3) is a well-known fact from linear algebra. So let us treat the equivalence of (1) and (2).

Let $P(x)$ be a monic polynomial of degree n with real coefficients. In the algebra $\mathbb{R}[x]$ of polynomials in x over $\mathbb{R}$ let I be the ideal generated by $P(x)$, and consider the algebra $T = \mathbb{R}[x]/I$. Now, $1, x, x^2, \dots, x^{n-1}$ are linearly independent in T, and x^n is a linear combination of them. Hence, $\dim T = \deg P(x) = n$. On T we have the trace map $T\to\mathbb{R}$, which is defined as usual: If $a\in T$, then a induces the multiplication $\bar{a}: T\to T$ with $b\mapsto ab$, and we put $\operatorname{trace}(a) := \operatorname{trace}(\bar{a})$. Then, $(a \mid b) := \operatorname{trace}(ab)$ is a symmetric bilinear form, and we can associate a quadratic form $F(a) := \operatorname{trace}(a^2)$.

Let J be the Jacobson radical of T, i.e., $J = \{a\in T : a^k = 0 \text{ for some } k\in\mathbb{N}\}$, and set $\bar{T} = T/J$. Then it is not hard to see that J is the kernel of the form F. Since each ideal in T is generated by a single element, we see that $\bar{T} = \mathbb{R}^{\oplus k}\oplus\mathbb{C}^{\oplus s}$, where k and $2s$ are the numbers of pairwise distinct real and complex roots of $P(x)$, respectively. By this identification, the class of the polynomial x maps to $\bar{x} = (\beta_1,\dots,\beta_k,\beta_{k+1},\dots,\beta_{k+s})$, where $\beta_1,\dots,\beta_k$ are the distinct real roots and

$\beta_{k+1}, \bar{\beta}_{k+1}, \ldots, \beta_{k+s}, \bar{\beta}_{k+s}$ the distinct complex roots of P. The trace map factors through $\bar{T}$ and gives

$$\operatorname{trace}(\lambda_1, \lambda_2, \ldots, \lambda_{k+s}) = \sum_{i=1}^{k} m_i \lambda_i + \sum_{j=1}^{s} m_{k+j}(\lambda_{k+j} + \bar{\lambda}_{k+j}),$$

where m_i is the multiplicity of the root β_i $(1 \le i \le k+s)$.

We assert that the quadratic form $F(a) = \operatorname{trace}(a^2)$ (considered as form on $\bar{T}$) is positive definite if and only if $s = 0$. This can be easily seen from the following formula

$$F(a) = \operatorname{trace}(a^2) = \operatorname{trace}(a_1^2, \ldots, a_{k+s}^2) = \sum_{i=1}^{k} m_i a_i^2 + \sum_{j=1}^{s} m_{k+j}(a_{k+j}^2 + \bar{a}_{k+j}^2).$$

Moreover, the signature of F is the number of distinct real roots of $P(x)$, namely k, since

$$\begin{aligned}
&F(\lambda_1, \ldots, \lambda_k, x_{k+1} + iy_{k+1}, \ldots, x_{k+s} + iy_{k+s}) \\
&= \operatorname{trace}(\lambda_1^2, \ldots, \lambda_k^2, x_{k+1}^2 - y_{k+1}^2 + 2ix_{k+1}y_{k+1}, \ldots, x_{k+s}^2 - y_{k+s}^2 + 2ix_{k+s}y_{k+s}) \\
&= \sum_{i=1}^{k} m_i \lambda_i^2 + \sum_{j=1}^{s} m_{k+j}(2x_{k+j}^2 - 2y_{k+j}^2).
\end{aligned}$$

Let us interpret what we have done so far. Since J is the kernel of the form F, we see that the rank of F equals $k + 2s$, that is the number of distinct roots of $P(x)$. If we consider the basis $1, \bar{x}, \ldots, \bar{x}^{n-1}$ of $\bar{T}$, we find immediately that the matrix of F in this basis is the Bezoutiant, and, therefore, the statements of the theorem follow by the considerations about F. □

3.1.3. LEMMA. *For $k \le n$ let A and B be matrices with k rows and n columns. The we have*

$$\det(AB^\top) = \sum_{1 \le i_1 < \cdots < i_k \le n} \det A_{i_1,\ldots,i_k} \cdot \det B_{i_1,\ldots,i_k}$$

where $A_{i_1,\ldots,i_k}$ is the minor of A with rows and columns $i_1, \ldots, i_k$.

PROOF. Let a_i and b_i, for $1 \le i \le k$, be the i-th row of A and B, respectively. So $a_i \in (\mathbb{R}^n)^*$ and $b_i^\top$. In terms of the standard basis e_i of $\mathbb{R}^n$ and its dual basis e^i of $(\mathbb{R}^n)^*$ we have

$$\begin{aligned}
&\det(AB^\top) = \det(\langle a_i, b_j^\top \rangle) = \langle a_1 \wedge \cdots \wedge a_k, b_1^\top \wedge \cdots \wedge b_k^\top \rangle \\
&= \left\langle \sum_{i_1 < \cdots < i_k} \det(\langle a_p, e_{i_q} \rangle) e^{i_1} \wedge \cdots \wedge e^{i_k}, \sum_{j_1 < \cdots < j_k} \det(\langle e^{j_q}, b_p^\top \rangle) e_{j_1} \wedge \cdots \wedge e_{j_k} \right\rangle \\
&= \sum_{i_1 < \cdots < i_k} \det(\langle a_p, e_{i_q} \rangle) \det(\langle b_p, e_{i_q} \rangle) \\
&= \sum_{i_1 < \cdots < i_k} \det A_{i_1,\ldots,i_k} \cdot \det B_{i_1,\ldots,i_k}.
\end{aligned}$$

□

3.1.4. Symmetric polynomials. Let $\mathbb{R}[x_1, \ldots, x_n]$ be the polynomial algebra in n variables over $\mathbb{R}$. A polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$ is called *symmetric* if $f(x_{\tau(1)}, \ldots, x_{\tau(n)}) = f(x_1, \ldots, x_n)$ for all $\tau \in \mathrm{S}_n$, the symmetric group on n elements. The subalgebra of symmetric polynomials is denoted by $\mathbb{R}[x_1, \ldots, x_n]^{\mathrm{S}_n}$. The elementary symmetric functions σ_i as well as the Newton polynomials s_i are symmetric.

THEOREM. *Every symmetric polynomial can be written as a polynomial in the elementary symmetric functions:*

$$\mathbb{R}[x_1,\ldots,x_n]^{\mathrm{S}_n} = \mathbb{R}[\sigma_1,\ldots,\sigma_n].$$

PROOF. Consider the lexicographic order on monomials:

$$x_1^{a_1}\cdots x_n^{a_n} < x_1^{b_1}\cdots x_n^{b_n}$$

if and only if the first non-zero difference $b_i - a_i$ is positive.

We proceed by induction over this order. Let $f \in \mathbb{R}[x_1,\ldots,x_n]^{\mathrm{S}_n}$. The action of S_n on $\mathbb{R}[x_1,\ldots,x_n]$ maps homogeneous polynomials to homogeneous polynomials of the same degree. Hence, a polynomial is symmetric if and only if each of its homogeneous components is symmetric, and so we may assume without loss that f is homogeneous. Let $x_1^{a_1}\cdots x_n^{a_n}$ be the largest monomial appearing with non-zero coefficient, say a, in f. Then $a_{i+1} \le a_i$ for all i. For contradiction suppose i_0 is the smallest i such that $a_{i+1} > a_i$. The transposition interchanging i_0 and i_0+1 belongs to S_n and since f is symmetric, the monomial $x_1^{a_1}\cdots x_{i_0}^{a_{i_0+1}} x_{i_0+1}^{a_{i_0}}\cdots x_n^{a_n}$ also appears in f with coefficient a. But

$$x_1^{a_1}\cdots x_n^{a_n} < x_1^{a_1}\cdots x_{i_0}^{a_{i_0+1}} x_{i_0+1}^{a_{i_0}}\cdots x_n^{a_n},$$

a contradiction. Then

$$\sigma_1^{a_1-a_2}\sigma_2^{a_2-a_3}\cdots\sigma_n^{a_n}$$

also contains $x_1^{a_1}\cdots x_n^{a_n}$ as largest monomial. It follows that

$$f - a\sigma_1^{a_1-a_2}\sigma_2^{a_2-a_3}\cdots\sigma_n^{a_n}$$

is a symmetric polynomial whose largest monomial is lower in the lexicographic order than the largest monomial of f. By repeating this process, we must eventually reach the zero polynomial. That means that we can express f as a polynomial in the elementary symmetric functions σ_i. □

3.2. Vandermonde varieties and Vandermonde functions

We follow [**Meg92**].

3.2.1. DEFINITION. Let us denote by $X_k(c)$ (respectively $X_k^{\mathbb{C}}(c)$) the *real* (respectively *complex*) *Vandermonde variety of order* k which is the algebraic subset of $\mathbb{R}^n$ (respectively $\mathbb{C}^n$) defined by the following k equations:

$$x_1^j + \cdots + x_n^j = c_j \qquad (1 \le j \le k).$$

We shall denote $s^n = (s_1,\ldots,s_n) : \mathbb{R}^n \to \mathbb{R}^n$, where $s_i(x) = \sum_{j=1}^n x_j^i$, as *Vandermonde mapping.*

REMARK. More generally one can study 'weighted' Vandermonde varieties given by

$$a_1x_1^j + \cdots + a_nx_n^j = c_j \qquad (1 \le j \le k).$$

where a real vector of non-zero weights $(a_1,\ldots,a_n)$ is fixed; [**Arn86**], [**Giv87**], [**Kos89**].

3.2.2. PROPOSITION. *A complex Vandermonde variety of order k has codimension equal to k.*

PROOF. Let $s_j = \sum_{i=1}^n x_i^j$ be the j-th Newton polynomial, and consider the map $s^n = (s_1,\ldots,s_n) : \mathbb{C}^n \to \mathbb{C}^n$. Then s^n is surjective and has finite fibers. Let Y be an $(n-k)$-dimensional linear subspace in $\mathbb{C}^n$ obtained by fixing the first k coordinates. The variety $X = (s^n)^{-1}(Y)$ is a Vandermonde variety of order k. The restriction of s^n to X is still a polynomial map with finite fibers and image Y, as is the restriction to any irreducible component X_i of X, i.e., $s^n|_{X_i} : X_i \to Y_i = s^n(X_i)$.

Since X_i and Y_i are both irreducible, in a non-empty open subset of Y_i the dimension of the fibers equals $\dim X_i - \dim Y_i$, whence $\dim X_i = \dim Y_i$. Moreover, $Y_i \subseteq Y$ implies $\dim Y_i \leq n-k$. So $\dim X_i \leq n-k$ for all i and thus $\dim X \leq n-k$. We have $\dim X = n-k$, since k polynomials define in $\mathbb{C}^n$ an algebraic subset of dimension at least $n-k$. □

3.2.3. Proposition. *If the real Vandermonde variety $X_k(c)$ contains a point having at least k distinct coordinates, then it has codimension k.*

Proof. It follows from proposition 3.2.2 that $\dim X_k(c) \leq n-k$. Consider the map $s^k = (s_1, \ldots, s_k) : \mathbb{R}^n \to \mathbb{R}^k$ where s_j is again the j-th Newton polynomial. Let $a = (a_1, \ldots, a_n)$ be a point in $\mathbb{R}^n$ with k distinct coordinates, and let $c = s^k(a)$, i.e., $a \in X_k(c)$. Then for the Jacobian matrix we get

$$ds^k(a) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 2a_1 & 2a_2 & \dots & 2a_n \\ \vdots & \vdots & \ddots & \vdots \\ ka_1^{k-1} & ka_2^{k-1} & \dots & ka_n^{k-1} \end{pmatrix}. \tag{3.4}$$

The minors of order k of $ds^k(a)$ are products of Vandermonde determinants by non-zero constants. Thus there exists one non-vanishing k-minor, namely the one involving the k distinct coordinates of a. Hence, $ds^k(a)$ has maximal rank k, and s^k is a submersion at a. It follows that the local dimension at a of the variety $X_k(c) = (s^k)^{-1}(c)$ equals $n-k$. Since the global dimension does not exceed $n-k$, it must be exactly $n-k$. □

3.2.4. Proposition. *Suppose that c is such that $X_k(c)$ has exactly codimension k. Then:*

(1) *The singular points of $X_k(c)$ are those points of $X_k(c)$ having less than k distinct coordinates.*
(2) *The critical points of s_{n+1} on the regular part $X_k(c)_{\mathrm{reg}}$ of $X_k(c)$ are those points having exactly k distinct coordinates.*

Proof. The first statement follows directly from (3.4).

To (2): Let $x = (x_1, \ldots, x_n)$ be a critical point of s_{k+1} on $X_k(c)_{\mathrm{reg}}$. Then there exist Lagrange multipliers $p_1, \ldots, p_k$ such that $ds_{k+1} = p_1 ds_1 + \cdots + p_k ds_k$ at x. If $L = s_{k+1} - p_1(s_1 - c_1) - \cdots - p_k(s_k - c_k)$ is the Lagrange function, then x solves the system of equations $dL = 0, s_1 = c_1, \ldots, s_k = c_k$. If we define $F(z) = (k+1)z^k - kp_k z^{k-1} - \cdots - p_1$, then $dL(x) = 0$ if and only if $F(x_i) = 0$ for all $1 \leq i \leq n$. It follows that x cannot have more than k distinct coordinates. Since x is regular, there are exactly k coordinates, by (1). Conversely, assume that x is a point of $X_k(c)$ having k distinct coordinates $x_1, \ldots, x_k$. Defining

$$p_{k+1-j} = (-1)^{j+1} \frac{k+1}{k+1-j} \sum_{1 \leq i_1 < \cdots < i_j \leq k} x_{i_1} \cdots x_{i_j},$$

for $1 \leq j \leq k$, gives k Lagrange multipliers for which x is a critical point. □

Remark. In the terminology of Arnol'd [**Arn86**] the function s_{k+1} restricted to the variety $X_k(c)$ is a *Vandermonde function.*

3.2.5. Proposition. *Fix a critical point $x = (x_1, \ldots, x_n)$ of s_{k+1} on $X_k(c)_{\mathrm{reg}}$ and let $z_1 > z_2 > \cdots > z_k$ be its k distinct coordinates. Suppose the coordinate z_i appears m_i times in $x = (x_1, \ldots, x_n)$. Let $r_i = m_i - 1$. Then the Hessian of s_{k+1} on $X_k(c)$ at the point x is the sum of a positive definite quadratic form on $\mathbb{R}^a$, where $a = r_1 + r_3 + \cdots$, and of a negative definite quadratic form on $\mathbb{R}^b$, where $b = r_2 + r_4 + \cdots$. In other words, s_{k+1} is a Morse function with Morse index b.*

PROOF. Without loss of generality we may assume that:

$$x = (\underbrace{z_1, \dots, z_1}_{r_1 \text{ times}}, \underbrace{z_2, \dots, z_2}_{r_2 \text{ times}}, \dots, \underbrace{z_k, \dots, z_k}_{r_k \text{ times}}, z_1, \dots, z_k).$$

We obtain a non-zero minor of $ds^k(x)$ of order k if we choose the last k columns, see (3.4). That means that the first $n-k$ coordinates may be used as a system of local coordinates for $X_k(c)$ in a neighborhood of x. We regard in a neighborhood of x the Lagrange function L from the proof of proposition 3.2.4 as a function of these $n-k$ first coordinates. Let $\tilde{x} = (x_1, \dots, x_{n-k})$ be the $(n-k)$-tuple of the first $n-k$ coordinates of the critical point x. Then we still have $dL(\tilde{x}) = 0$. We find $\frac{\partial^2 L}{\partial x_i \partial x_j} = 0$ for $i \neq j$ and $\frac{\partial^2 L}{\partial x_i^2} = F'(x_i)$, where F is the polynomial defined in the proof of proposition 3.2.4. It follows that

$$d^2L(\tilde{x})(h_1, \dots, h_{n-k}) = (h_1^2 + \dots + h_{r_1}^2)F'(z_1) + \dots + (h_{n-k-r_k}^2 + \dots + h_{n-k}^2)F'(z_k).$$

Since $F'(z_1) > 0$, $F'(z_2) < 0$, $F'(z_3) > 0$, and so on, the statement of the proposition is proved. $\square$

COROLLARY. *A critical point x is a strict local minimum of s_{k+1} on $X_k(c)$ if the vector $(m_1, \dots, m_k)$ of the multiplicities of the distinct coordinates $z_1 > \dots > z_k$ in x has the form $(r, 1, s, 1, \dots)$ and is a strict local maximum if it has the form $(1, r, 1, s, 1, \dots)$.* $\square$

3.2.6. LEMMA. *The reconstruction of the topology of a level set of a function f on $\mathbb{R}_+^a \times \mathbb{R}_+^b$ in the neighborhood of the critical point $(0,0)$ with non-degenerate Hessian $F = Q_+ + Q_-$ is trivial if $a, b > 0$ and consists of the birth (death) of a simplex otherwise. (Here we write $\mathbb{R}_+$ for $[0, \infty)$.)*

PROOF. Let us first verify the lemma with F in place of f. We may suppose that $F(x,y) = x_1^2 + \dots + x_a^2 - (y_1^2 + \dots + y_b^2)$. Let $\epsilon > 0$ and consider

$$F^{-1}(\epsilon) = \{(x,y) \in \mathbb{R}_+^a \times \mathbb{R}_+^b : \textstyle\sum_{i=1}^a x_i^2 = \sum_{i=1}^b y_i^2 + \epsilon\}.$$

Assume that a and b are positive. Then the mapping

$$\begin{aligned} F^{-1}(\epsilon) &\longrightarrow \Delta^{a-1} \times \mathbb{R}_+^b \\ (x,y) &\longmapsto \left(\frac{x}{\|x\|}, y\right) \end{aligned}$$

is a homeomorphism with inverse

$$\begin{aligned} \Delta^{a-1} \times \mathbb{R}_+^b &\longrightarrow F^{-1}(\epsilon) \\ (x,y) &\longmapsto \left(x\sqrt{\|y\|^2 + \epsilon}, y\right), \end{aligned}$$

where $\|.\|$ is the Euclidean norm, and $\Delta^{a-1} = \{x \in \mathbb{R}_+^a : \|x\| = 1\}$. Since the simplex Δ^{a-1} is homeomorphic to $[0,1]^{a-1}$, $F^{-1}(\epsilon)$ is homeomorphic to $[0,1]^{a-1} \times \mathbb{R}_+^b$, and to $\mathbb{R}_+^{a+b-1}$, since $[0,1] \times \mathbb{R}_+$ is homeomorphic to $\mathbb{R}_+^2$. Analogously we find that $F^{-1}(-\epsilon)$ is homeomorphic to $\mathbb{R}_+^{a+b-1}$. Now,

$$F^{-1}(0) = \{(x,y) \in \mathbb{R}_+^a \times \mathbb{R}_+^b : \textstyle\sum_{i=1}^a x_i^2 = \sum_{i=1}^b y_i^2\}$$

is the cone over $\Delta^{a-1} \times \Delta^{b-1}$ obtained by contracting the base of the cylinder $\Delta^{a-1} \times \Delta^{b-1} \times \mathbb{R}_+$. Thus, if ϕ is a homeomorphism between $\Delta^{a-1} \times \Delta^{b-1}$ and Δ^{a+b-2}, then

$$\begin{array}{ccccc} \Delta^{a+b-2} \times \mathbb{R}_+ & \longleftarrow & \Delta^{a-1} \times \Delta^{b-1} \times \mathbb{R}_+ & \longrightarrow & F^{-1}(0) \\ (\phi(x,y), t) & \longleftarrow & (x, y, t) & \longrightarrow & (tx, ty) \end{array}$$

shows that $F^{-1}(0)$ may be obtained by contracting the base of the cylinder $\Delta^{a+b-2} \times \mathbb{R}_+$. Therefore, $F^{-1}(0)$ is homeomorphic to the cone over Δ^{a+b-2}, i.e., to $\mathbb{R}_+^{a+b-1}$.

Now suppose that a or b is zero. Say $b = 0$. Then we have $F^{-1}(-\epsilon) = \emptyset$, $F^{-1}(0) = \{0\}$, and $F^{-1}(\epsilon) = \Delta^{a-1}$, which is the birth of a simplex. For $a = 0$ we get the death of a simplex.

We complete the proof by considering $f = F + o(\|x^2\|)$. We claim the following homeomorphisms:

$$F^{-1}(-\epsilon) \cong f^{-1}(-\epsilon), \quad F^{-1}(0) \cong f^{-1}(0), \quad \text{and} \quad F^{-1}(\epsilon) \cong f^{-1}(\epsilon).$$

There is no difficulty for non-critical level sets. So consider the critical level set $V = f^{-1}(0)$ and the following blow-up:

$$\begin{aligned} \pi : \Delta^{a+b-1} \times \mathbb{R}_+ &\longrightarrow \mathbb{R}_+^a \times \mathbb{R}_+^b \\ (x, y, t) &\longmapsto (tx, ty). \end{aligned}$$

Then $\pi^{-1}(V)$ is defined by the equation $F(tx, ty)+$ rest $= 0$, which can be rewritten as $t^2(Q_+(x) + Q_-(y)) + t^3\text{rest} = 0$. We may decompose $\pi^{-1}(V)$ into

$$\tilde{V} = \{(x, y, t) \in \Delta^{a+b-1} \times \mathbb{R}_+ : Q_+(x) + Q_-(y) + t \text{ rest } = 0\}$$

and the exceptional divisor

$$ED = \Delta^{a+b-1} \times \{0\}.$$

Since not all x_i and y_i are zero, for t small enough the variety $\tilde{V}$ given by the equation

$$\sum_{i=1}^{a} x_i^2 = \sum_{i=1}^{b} y_i^2 - t \text{ rest}$$

has the same topological type as the variety associated to

$$\sum_{i=1}^{a} x_i^2 = \sum_{i=1}^{b} y_i^2.$$

Hence in a neighborhood of $t = 0$ the variety $\tilde{V}$ is a cylinder over

$$\{(x, y) \in \mathbb{R}_+^{a+b} : \textstyle\sum_{i=1}^a x_i^2 = \sum_{i=1}^b y_i^2 \text{ and } \sum_{i=1}^a x_i^2 + \sum_{i=1}^b y_i^2 = 1\},$$

that is, over $\Delta^{a-1} \times \Delta^{b-1}$. Then V is obtained by contracting to a point the intersection $\tilde{V} \cap ED = \Delta^{a-1} \times \Delta^{b-1}$. Consequently, V is the cone over $\Delta^{a-1} \times \Delta^{b-1}$ which shows that $F^{-1}(0) \cong f^{-1}(0)$.

This completes the proof. $\square$

3.2.7. Theorem. *Let $K = \{x \in \mathbb{R}^n : x_1 \leq \cdots \leq x_n\}$, $0 < k < n$, $c' = (c_1, \ldots, c_{k+1}) \in \mathbb{R}^{k+1}$, and $c = (c_1, \ldots, c_k)$. If $X_k(c)$ is non-singular and if $K \cap X_k(c)$ is connected, then $K \cap X_{k+1}(c')$ is connected as well.*

Proof. The chamber K is naturally Whitney stratified, see 9.6.4 and 9.6.5. The stratum K_l of dimension l is the set of points in K with exactly l distinct coordinates. Let us suppose that $X_k(c)$ is non-singular.

Claim (1). *$X_k(c)$ is transversal to all strata of K.*

We write $z = (z_1, \ldots, z_l)$ for a point in $K_l \cap X_k(c)$ where $z_1 < \cdots < z_l$. The tangent space at z to $X_k(c)$ is the set of vectors $x = (x_1, \ldots, x_n)$ such that

$$\begin{cases} x_1 + \cdots + x_n = 0 \\ z_1 x_1 + \cdots + z_l x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ z_1^{k-1} x_1 + \cdots + z_l^{k-1} x_n = 0. \end{cases}$$

We have to show that this $(n-k)$-dimensional space is transversal to the stratum K_l. Since $X_k(c)$ is non-singular it does not intersect strata of dimension lower than k, by proposition 3.2.4. So we assume $l \geq k$. Let $m = (m_1, \dots, m_l)$ be the vector of multiplicities m_i of z_i in z. A typical vector y of the tangent space $T_z K_l$ is:

$$y = (\underbrace{y_1, \dots, y_1}_{m_1 \text{ times}}, \underbrace{y_2, \dots, y_2}_{m_2 \text{ times}}, \dots, \underbrace{y_l, \dots, y_l}_{m_l \text{ times}}).$$

If $y \in T_z X_k(c)$, then y solves the following system of equations:

$$\begin{cases} m_1 y_1 + \cdots + m_l y_l = 0 \\ m_1 z_1 y_1 + \cdots + m_l z_l y_l = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ m_1 z_1^{k-1} y_1 + \cdots + m_l z_l^{k-1} y_l = 0. \end{cases}$$

It follows that the intersection $T_z K_l \cap T_z X_k(c)$ has dimension $l-k$, which is the minimal dimension of the intersection of a space with codimension k and a space of dimension l. Therefore, $X_k(c)$ is transversal to K_l. This shows claim 1.

It follows that the stratification of $K \cap X_k(c)$, whose strata are the intersections of $X_k(c)$ with strata of K, is a Whitney stratification; compare with 9.6.5.

The function $s_{k+1} = \sum_{i=1}^n x_i^{k+1}$ is smooth on $\mathbb{R}^n$. Let us consider the restriction of s_{k+1} to $X = K \cap X_k(c)$, and consider $X_{\leq a} = \{x \in X : s_{k+1}(x) \leq a\}$. Then the topological type of $X_{\leq a}$ remains constant as a varies within the open interval between two adjacent critical values of s_{k+1}. A critical point of s_{k+1} is a critical point of any restriction of s_{k+1} to a stratum of X.

Claim (2). *The critical points of $s_{k+1}|_X$ are exactly the isolated points of the 0-dimensional stratum $K_k \cap X_k(c)$ of X.*

The isolated points of the 0-dimensional stratum $K_k \cap X_k(c)$ of X are those points of X having exactly k distinct coordinates and hence are critical points of $s_{k+1}|_X$. Now we consider the restriction of s_{k+1} to any stratum of X of higher dimension. More precisely, let us study the restriction of s_{k+1} to any component Y of a stratum $K_l \cap X_k(c)$ with $l > k$, where this component is characterized by the multiplicity vector $m = (m_1, \dots, m_l)$ of its points. It is equivalent to the study of the function $\hat{s}_{k+1} = \sum_{i=1}^l m_i x_i^{k+1}$ on the subvariety of $\mathbb{R}^l$ given by:

$$m_1 x_1^j + \cdots + m_l x_l^j = c_j \quad (1 \leq j \leq k) \qquad \text{and} \qquad x_1 < \cdots < x_l.$$

This study was made before with weights m_i all equal to 1 and remains valid in the case of any positive weights. By proposition 3.2.4, the critical points of $\hat{s}_{k+1} = \sum_{i=1}^l m_i x_i^{k+1}$ on this subvariety are the points having k distinct coordinates. But there are no such points, as we fixed $k < l$. This shows claim 2.

Let $z = (z_1, \dots, z_k) \in \mathbb{R}^n$ be a point in $K \cap X_k(c)$ having exactly k distinct coordinates. Then z is a critical point of s_{k+1} on $X_k(c)$, and, if $m = (m_1, \dots, m_k)$ is the multiplicity vector of z, the Hessian of s_{k+1} at z is a quadratic form of signature (a, b), where $a = \sum(m_{2i-1} - 1)$ and $b = \sum(m_{2i} - 1)$; see proposition 3.2.5. Let $\phi : \mathbb{R}^n \to \mathbb{R}^n$ be defined by

$$\phi(x_1, \dots, x_n) = (x_2 - x_1, \dots, x_{m_1} - x_{m_1 - 1}, x_{m_1+2} - x_{m_1+1}, \dots, x_n - x_{n-1}, s_1, \dots, s_k).$$

It is easily seen that ϕ is a local diffeomorphism at z, $\phi(z) = (0, \dots, 0, c_1, \dots, c_k)$, and ϕ sends a neighborhood of z in $K \cap X_k(c)$ onto a neighborhood of $(0, \dots, 0, c_1, \dots, c_k)$ in $\mathbb{R}_+^{a+b} \times \{(c_1, \dots, c_k)\}$. This shows that we may reduce the study of s_{k+1} on $K \cap X_k(c)$ in a neighborhood of z to the study of a function f on $\mathbb{R}_+^{a+b}$ in a neighborhood of the origin, which is a critical point of f, admitting a Hessian having the same signature. But this study has been done in lemma 3.2.6.

Reconstructing $K \cap X_k(c)$ level by level, we have seen that the topological type does not change between two consecutive critical values, and at a critical point a simplex may appear or disappear. Suppose the reconstructed variety is connected. Then:

CLAIM (3). *There are only two critical points on $K \cap X_k(c)$. At the first critical points a simplex appears, and it disappears at the second critical point.*

If two simplices appear at two distinct critical points, then there exists another critical point where the two simplices are joined, since $K \cap X_k(c)$ is connected. By lemma 3.2.6, such a critical point (it is a saddle) is not possible. This implies the claim.

By claim 3, the level sets of s_{k+1} on $K \cap X_k(c)$ are all contractible. Since these level sets are exactly the varieties $K \cap X_{k+1}(c')$, the theorem is proved. $\square$

3.2.8. COROLLARY. *For almost all $c \in \mathbb{R}^n$, the set $K \cap X_k(c)$ is either empty or connected.*

PROOF. Using induction on k it follows from theorem 3.2.7 and its proof that, for almost all c, $K \cap X_k(c)$ is either empty or connected, since it is obvious for $k = 1, 2$. $\square$

REMARK. It is shown in [**Kos89**] that actually for all $c \in \mathbb{R}^n$, the set $K \cap X_k(c)$ is either empty or contractible.

COROLLARY. *Let $0 < k < n$. Let us denote by $p_k(a)$ the property that $X_k(a)$ is empty or $X_k(a)$ is non-singular and $K \cap X_k(a)$ is connected. Then $p_k(a)$ is valid for almost all $a \in \mathbb{R}^k$. More precisely, it is true outside a hypersurface of $\mathbb{R}^k$.*

PROOF. Let us use induction on k. The hyperplane $X_1(a)$ is non-singular and $K \cap X_1(a)$ is connected. Suppose that $p_k(a)$ is true for almost all $a \in \mathbb{R}^k$, more precisely, it holds outside of a hypersurface H of $\mathbb{R}^k$. Now, let $a \in \mathbb{R}^{k+1}$. If $a \notin H \times \mathbb{R}$, then theorem 3.2.7 implies that $K \cap X_{k+1}(a)$ is connected. Recall that $X_{k+1}(a) = (s^{k+1})^{-1}(a)$ where $s^{k+1} = (s_1, \dots, s_{k+1}) : \mathbb{R}^n \to \mathbb{R}^{k+1}$. According to Sard's theorem, the critical values of s^{k+1} lie on a hypersurface H' of $\mathbb{R}^{k+1}$. If $a \notin H'$, then either $(s^{k+1})^{-1}(a)$ is empty or s^{k+1} is a submersion at any point of $(s^{k+1})^{-1}(a)$, and $X_{k+1}(a) = (s^{k+1})^{-1}(a)$ is a non-singular variety of codimension $k+1$. Consequently, if $a \notin H' \cup H \times \mathbb{R}$, which is a hypersurface of $\mathbb{R}^{k+1}$, then $p_{k+1}(a)$ holds true. This completes the induction and hence the proof. $\square$

3.3. Maximal hyperbolic polynomials

3.3.1. DEFINITION. Let n and s be two integers such that $0 \leq s < n$. We will say that a monic hyperbolic polynomial $P(x) = x^n + a_1 x^{n-1} + \cdots + a_n$ is *locally s-maximal* (respectively *locally s-minimal*) if there exists a neighborhood of P (identified with $(a_1, a_2, \dots, a_n) \in \mathbb{R}^n$) where no polynomial $P(x) + c_0 x^s + c_1 x^{s-1} + \cdots + c_s$ with $c_0 > 0$ (respectively $c_0 < 0$) is hyperbolic.

The hyperbolic polynomial P is said to be *globally s-maximal* (respectively *globally s-minimal*) if $P(x) + c_0 x^s + c_1 x^{s-1} + \cdots + c_s$ with $c_0 > 0$ (respectively $c_0 < 0$) is never hyperbolic.

REMARK. Geometrically: Finding a polynomial $Q(x) = c_0 x^s + c_1 x^{s-1} + \cdots + c_s$ such that $P + Q$ is hyperbolic is equivalent to finding a curve $y = -Q(x)$ which intersects the curve $y = P(x)$ exactly n times.

EXAMPLE. x^2 is 0-maximal, x^3 and $x(x^2-1)^2$ are 1-minimal. This can be easily seen geometrically as indicated in above remark, or formally with theorem 3.3.6.

3.3.2. Definition. . If P is a monic hyperbolic polynomial, then it may be written uniquely as

$$P(x) = (x - x_1)^{m_1} \cdots (x - x_k)^{m_k} \quad \text{with} \quad x_1 > x_2 > \cdots > x_k.$$

The k-tuple $(m_1, \ldots, m_k)$ is called the *multiplicity vector* of P.

To any monic hyperbolic polynomial P we may associate an integer s_P which depends only on the multiplicity vector of P, namely:

(1) If P has only simple roots, put $s_P = -1$.
(2) Otherwise, let $s_P = l + \sum_{m_i \geq 2}(m_i - 2)$ where l is the number of the odd sequences of consecutive 1 in the multiplicity vector of P considering only sequences which are between two multiplicities greater or equal to 2.

We shall say that P is *even* when its multiplicity vector begins with an even number of consecutive 1, and *odd* otherwise.

The hyperbolic polynomial P is said to be *alternate-even* if $m_{2i} = 1$ for all i and *alternate-odd* if $m_{2i+1} = 1$ for all i.

Example. Let P have the multiplicity vector $(1, 1, 1, 4, 1, 1, 3, 1, 5, 1, 1)$. Then $s_P = 1 + (4 - 2) + (3 - 2) + (5 - 2) = 7$, and P is odd. Further, P is neither alternate-even nor alternate-odd.

3.3.3. Proposition. *An alternate-even (respectively alternate-odd) hyperbolic polynomial P is locally s_P-maximal (respectively locally s_P-minimal).*

Proof. Let P be an alternate-even hyperbolic polynomial having exactly $k < n$ distinct roots (if P has only simple roots there is nothing to prove). Let $X_k(c)$ be the real Vandermonde variety of order k which contains the n-tuple $x = (x_1, \ldots, x_n)$ of roots of P. This n-tuple has the multiplicity vector $(r, 1, s, 1, \ldots)$ and is therefore a strict local minimum of s_{n+1} on $X_k(c)$, by corollary 3.2.5. Suppose Q is a hyperbolic polynomial of degree n the first k Newton polynomials of whose roots agree with those of P, we write $s_i(P) = s_i(Q)$ for $1 \leq i \leq k$. That means that the n-tuple consisting of the roots of Q lies on $X_k(c)$. If we write $\sigma_i(P)$ and $\sigma_i(Q)$ for the elementary symmetric functions in the roots of P and Q, then (3.1) implies that $\sigma_i(P) = \sigma_i(Q)$ for $1 \leq i \leq k$. If the roots of Q are close enough to those of P, we have $s_{k+1}(P) \leq s_{k+1}(Q)$, and, consequently, for the $(k+1)$-st coefficients in P and Q we find $(-1)^{k+1}\sigma_{k+1}(P) \geq (-1)^{k+1}\sigma_{k+1}(Q)$. It follows that P is locally $(n - k - 1)$-maximal. It is easy to see that $s_P = n - k - 1$. Analogously one shows that an alternate-odd hyperbolic polynomial P is locally s_P-minimal. $\square$

3.3.4. Proposition. *An alternate-even (respectively alternate-odd) hyperbolic polynomial P is globally s_P-maximal (respectively locally s_P-minimal).*

Proof. Let P be an alternate-even hyperbolic polynomial, let $x = (x_1, \ldots, x_k)$ be the k-tuple of its distinct roots, and let $(m_1, 1, m_3, 1, \ldots)$ be its multiplicity vector. Consider the following map $a : \mathbb{R}^k \to \mathbb{R}^k$ given by

$$(y_1, \ldots, y_k) \longmapsto (m_1 y_1 + y_2 + \cdots, \ldots, m_1 y_1^k + y_2^k + \cdots)$$

which is a local diffeomorphism at x. Thus, there exists a neighborhood U of x which is diffeomorphic to a neighborhood V of $a(x)$. We apply corollary 3.2: There exists a hypersurface H such that property $p_k(a')$ is true for all $a' \in V \backslash H$. By proposition 3.3.3, P is locally s_P-maximal. Note that $s_P = n - k - 1$. For contradiction assume that P is not globally s_P-maximal, i.e., there exists a polynomial $R = P + cx^{s_P} + \cdots$ which is hyperbolic, where c is positive. Let us consider a neighborhood V' of R and the following maps

$$\begin{array}{ccccc} V' & \xrightarrow{\pi} & \mathbb{R}^k & \xrightarrow{\psi} & \mathbb{R}^k \\ R' & \longmapsto & (\sigma_1, \ldots, \sigma_k) & \longmapsto & (s_1, \ldots, s_k) \end{array}$$

where σ_i is the i-th elementary symmetric function of the roots of R', and ψ is the diffeomorphism given by (3.1). We have $\psi(\pi(R)) = a(x)$. If Hyp_n denotes the set of monic hyperbolic polynomials of degree n (viewed as subset of $\mathbb{R}^n$), then $V' \cap \mathrm{Hyp}_n$ is a semialgebraic set of dimension n containing R. It follows that $W := \psi(\pi(V' \cap \mathrm{Hyp}_n))$ is a semialgebraic set of dimension k containing $a(x)$. Necessarily we have that $W \cap (V \backslash H) \neq \emptyset$.

Let $a' \in W \cap (V\backslash H) \neq \emptyset$ and $x' := a^{-1}(a')$. We associate to x' an alternate-even polynomial Q with the same multiplicity vector as P. To the polynomial Q we associate the n-tuple $\bar{x}'$ of its roots in increasing order. Then $\bar{x}' \in K \cap X_k(a')$ is a local minimum of the restriction $h := s_{k+1}|_{K\cap X_k(a')}$ of the $(k+1)$-st Newton sum s_{k+1}, by corollary 3.2.5; here K and $X_k(a')$ are defined as in the previous section. Since property $p_k(a')$ is valid and $X_k(a')$ is not empty, we conclude that $X_k(a')$ is non-singular and $K \cap X_k(a')$ is connected. By theorem 3.2.7, for all $a'' \in \mathbb{R}$, $K \cap X_{k+1}(a', a'')$ is connected as well. Let us apply this result to $a'' = h(\bar{x}')$. Thus, $K \cap X_{k+1}(a', h(\bar{x}')) = h^{-1}(h(\bar{x}'))$ is connected and contains an isolated point $\bar{x}'$ which is a strict local minimum of h. It follows that $h^{-1}(h(\bar{x}')) = \{\bar{x}'\}$. Since $K \cap X_k(a')$ is connected, we find that $\bar{x}'$ is a global minimum of h. Then Q is globally $s_P(= s_Q)$-maximal, by a similar reasoning as at the end of the proof of proposition 3.3.3.

Since $a' \in W$, there exists a hyperbolic polynomial R' in V' whose first k coefficients equal those of Q; namely $\psi(\pi(R')) = a'$. Now, R' is close to R, and Q is close to P. Thus, $R' - Q = c'x^{s_P} + \cdots$ with c' close to $c > 0$. If we choose V' to be sufficiently small, we may assume $c' > 0$. But $R' = Q + c'x^{s_P} + \cdots$ is hyperbolic, which contradicts that Q is globally s_P-maximal. It follows that P is globally s_P-maximal. □

3.3.5. Lemma. *Let $a_1, \dots, a_q$ be real numbers, and let P be a hyperbolic polynomial. If the polynomial $P(x) \cdot (x - a_1) \cdots (x - a_q)$ is globally $(s + q)$-maximal (respectively minimal), then P is globally s-maximal (respectively minimal).*

Proof. It is evident that if $P + Q$ is hyperbolic, then $P(x) \cdot (x - a_1) \cdots (x - a_q) + Q(x) \cdot (x - a_1) \cdots (x - a_q)$ is hyperbolic as well. Consequently, if P is not globally s-maximal (respectively minimal), then $P(x) \cdot (x - a_1) \cdots (x - a_q)$ is not globally $(s + q)$-maximal (respectively minimal). □

3.3.6. The escape from the space of hyperbolic polynomials.

Theorem. *Let P be a hyperbolic polynomial. Then we have:*
(1) *P is locally s-maximal (respectively minimal) if and only if P is globally s-maximal (respectively minimal).*
(2) *If $0 \leq s < s_P$, then P is s-maximal and s-minimal.*
(3) *If $s = s_P$ and P is even, then P is s_P-maximal but not s_P-minimal.*
(4) *If $s = s_P$ and P is odd, then P is s_P-minimal but not s_P-maximal.*
(5) *If $s > s_P$, then P is neither s-maximal nor s-minimal.*

Proof. We start with the following assertion.

Claim (1). *Let P be an even (respectively odd) hyperbolic polynomial. Then:*
- *If $s > s_P$, then P is neither locally s-maximal nor locally s-minimal.*
- *If $s = s_P$, then P is not locally s_P-minimal (respectively maximal).*

We shall show claim 1 first for hyperbolic polynomials P with $s_P = 0$. This condition implies that P has only simple or double roots and that its multiplicity vector contains only even sequences of consecutive simple roots; moreover, not all roots of P can be simple. If P is even, then its double roots are local minima of the curve $y = P(x)$, and all local maxima of this curve are positive. Let $\eta > 0$ denote the minimum of these local maxima. Then, for $\epsilon \in [0, \eta]$, the line $y = \epsilon$ intersects the curve $n = \deg P$ times (with multiplicities). This shows that P is not locally 0-minimal. Let us form the set of all roots of P and the abscissas of all intersection points of the curve $y = P(x)$ and the line $y = \eta$. We denote by a and b the minimum and the maximum of this set. Let R be a non-constant polynomial. There exist constants $m < M$ such that $m \le R(x) \le M$ for all $x \in [a, b]$. For $Q(x) := \epsilon\frac{R(x)-m}{M-m}$, where $0 < \epsilon < \eta$, we find $0 \le Q(x) \le \eta$ for all $x \in [a, b]$. It follows that the two curves $y = P(x)$ and $y = Q(x)$ intersect n times (with multiplicities). Hence, for $s \ge 1$, P is neither locally s-maximal nor locally s-minimal. For odd polynomials the arguments are similar. This proves claim 1 for the case $s_P = 0$. Now let as prove it for arbitrary hyperbolic polynomials P. Let $(m_1, \dots, m_p)$ be the multiplicity vector of $P(x) = \prod_{i=1}^{p}(x - x_i)^{m_i}$ and $s_P = l + \sum_{m_i \ge 2}(m_i - 2)$, as defined in 3.3.2. In each of the l odd sequences, we choose a simple root x_{j_k} $(1 \le k \le l)$. We define the hyperbolic polynomial $Q(x) = \prod_{m_i \ge 2}(x - x_i)^{m_i - 2} \prod_{k=1}^{l}(x - x_{j_k})$. We have $\deg Q = s_P$, Q divides P, and, if P_1 denotes the quotient, then $s_{P_1} = 0$. Using the fact that if $P_1 + R$ is hyperbolic then $Q(P_1 + R) = P + QR$ is hyperbolic, and applying claim 1 to P_1, yields that claim 1 holds for P as well.

Claim (2). *Let P be an even (respectively odd) hyperbolic polynomial. Then:*
- *If $0 \le s < s_P$, then P is globally s-maximal and globally s-minimal.*
- *If $s = s_P$, then P is globally s_P-maximal (respectively minimal).*

Firstly, we will prove claim 2 for alternate-even and alternate-odd polynomials. We may assume that not all roots of P are simple. Then an alternate-even (respectively alternate-odd) polynomial is even (respectively odd). Let P be alternate-even with multiplicity vector $(r, 1, s, 1, \dots)$. We know from proposition 3.3.4 that P is globally s_P-maximal. It remains to show that P is globally s-maximal and s-minimal for $0 \le s < s_P$. Let q be a positive integer such that $s + q = s_P$. We form the polynomial $Q_1(x) = P(x) \cdot (x - a_1) \cdots (x - a_q)$ for real numbers $a_q < \cdots < a_1 < \min\{\text{roots of } P\}$. This process consists in adding some 1 to the end of the multiplicity vector of P in order to transform it into the multiplicity vector of Q_1. Thus, Q_1 is alternate-even, and $s_{Q_1} = s_P$. By proposition 3.3.4, Q_1 is globally s_P-maximal, and hence P is globally s-maximal, by lemma 3.3.5. We form Q_2, keeping $a_2, \dots, a_q$ as above but choosing $a_1 > \max\{\text{roots of } P\}$. Then Q_2 is alternate-odd. Applying as before proposition 3.3.4, we may conclude that P is globally s-minimal. For alternate-odd polynomials P the proof is similar. Next we prove claim 2 for arbitrary hyperbolic polynomials. Let $P(x) = (x - x_1)^{m_1} \cdots (x - x_p)^{m_p}$ with $x_1 > x_2 > \cdots > x_p$. In the multiplicity vector of P we may replace each even sequence of consecutive 1 lying between two multiplicities greater than or equal to 2 by an odd sequence multiplying P by $(x - a_k)$, where a_k is not a root of P. After a finite number of such multiplications one obtains a polynomial $Q(x) = P(x) \cdot (x - a_1) \cdots (x - a_q)$ with $s_Q = s_P + q$ which is alternate-even (respectively alternate-odd) if P is even (respectively odd). Applying claim 2 to Q and using lemma 3.3.5, establishes claim 2 for P.

The statements in the theorem are immediate consequences of claim 1 and claim 2. The proof is complete. □

3.4. The space of hyperbolic polynomials is Whitney regular

3.4.1. Definition. A compact, finitely arc-wise connected subset $X \subseteq \mathbb{R}^n$ is called *Whitney regular*, if any two points x and y in X can be connected by a piecewise smooth curve γ in X such that for the length $l(\gamma)$ of γ we have

$$l(\gamma) \le C\|x - y\|,$$

where $C \geq 1$ is a constant not depending on the choice of x and y.

3.4.2. We introduce the following notation: Let Hyp_n denote the space of monic hyperbolic polynomials

$$P(x) = x^n - a_1 x^{n-1} + \cdots + (-1)^n a_n$$

of a fixed degree n (viewed as subspace of $\mathbb{R}^n$) and by Hyp_n^0 its intersection with $\{a_1 = 0, |a_2| \leq 1\}$. Let $K = \{x \in \mathbb{R}^n : x_1 \leq x_2 \leq \cdots \leq x_n\}$, $s^k = (s_1, \dots, s_k) : K \to \mathbb{R}^k$ where $s_i = \sum_{j=1}^n x_j^i$, and $\Pi^k = s^k(K)$. Moreover, let $K_0 = K \cap \{x \in \mathbb{R}^n : s_1(x) = 0, s_2(x) \leq 2\}$ and $s^k(K_0) = \Pi_0^k$. Then $\Pi^n \cong \mathrm{Hyp}_n$ and $\Pi_0^n \cong \mathrm{Hyp}_n^0$.

3.4.3. We will need the following corollary of remark 3.2.7:

Corollary. *We have:*

(1) *The mapping $s^n : K \to \Pi^n$ is a homeomorphism.*
(2) *The mapping s^k is a homeomorphism of the closure of every k-dimensional stratum of K onto its image.*
(3) *Every set of the kind $\Pi^n \cap \{y_j = c_j, c_j \in \mathbb{R}, 1 \leq j \leq k, k \leq n\}$ is either contractible or empty.*

Proof. To (1): The contractibility of the non-empty set $K \cap X_n(c)$, see remark 3.2.7, guarantees injectivity of the map $s^n : K \to \Pi^n$. Surjectivity is trivial. Since it is continuous and the spaces involved are Hausdorff, it is a homeomorphism on compact subsets of K onto its image. For every bounded set $\omega \subseteq \Pi^n$ there exists some constant $C > 0$ such that $(s^n)^{-1}(\omega) \subseteq \{x \in K : s_2(x) \leq C\}$, which is compact. Hence, $(s^n)^{-1}(\omega)$ is homeomorphically mapped onto ω. This implies (1).

To (2): The restriction of s^k to any k-dimensional stratum of K is a weighted k-dimensional Vandermonde mapping, for which the whole discussion is true as well. Statement (2) thus follows from (1).

To (3): This is an immediate consequence of remark 3.2.7 and (1). □

3.4.4. Proposition. *The image of the closure $\overline{S}$ of every k-dimensional stratum S of K under s^l for $k \leq l \leq n$ is a stratified manifold. It is the graph of an $(l-k)$-dimensional vector function defined on $s^k(\overline{S})$. The only k-dimensional stratum of this manifold is $s^l(S)$. The tangent bundle of $s^l(S)$ is continuously extended on the strata of $s^l(\overline{S})$ of non-maximal dimension. This extension is everywhere transversal to the subspace spanned by the last $(l-k)$ coordinates.*

Proof. By corollary 3.4.3, any k-dimensional stratum of K is mapped by s^k onto its image homeomorphically. Therefore we may consider the coordinates $y_{k+1}, \dots, y_n$ of the points in the image of a k-dimensional stratum S under s^n as functions of the coordinates $y_1, \dots, y_k$.

Let us calculate the partial derivatives $\frac{\partial y_l}{\partial y_s}$, for $k+1 \leq l \leq n$, $1 \leq s \leq k$, bearing in mind that $y_j = s_j(x)$. By changing the numeration of the coordinate axes we may achieve that the coordinates on the stratum S are $x_1, \dots, x_k$. Let $m = (m_1, \dots, m_k)$ denote the multiplicity vector of $(x_1, \dots, x_k)$. Then we have

$$\frac{\partial y_l}{\partial y_s} = \sum_{i=1}^k \frac{\partial y_l}{\partial x_i}\frac{\partial x_i}{\partial y_s} = \sum_{i=1}^k l m_i x_i^{l-1}\frac{\partial x_i}{\partial y_s} = \frac{l}{w}\left(\sum_{i=1}^k m_i x_i^{l-1} A_{si}\right), \tag{3.5}$$

where $w = \det\left(\frac{\partial y_j}{\partial x_t}\right)_{jt} = b\prod_{q<r}(x_q - x_r)$, $b \neq 0$, and A_{si} is the cofactor of the element $\frac{\partial y_s}{\partial x_i}$ in the matrix $\left(\frac{\partial y_j}{\partial x_t}\right)_{jt}$.

Let us put $x_\mu = x_\nu$ in (3.5). For $i \neq \mu, \nu$ we have $A_{si} = 0$ (two proportional columns). Moreover, we have $m_\mu A_{s\mu} + m_\nu A_{s\nu} = 0$. Thus, the numerator in the right-hand side of (3.5) vanishes when $x_\mu = x_\nu$, i.e., it is represented in the form $wf(x_1, \dots, x_k)$, where f is a polynomial. Consequently,

$$\frac{\partial y_l}{\partial y_s} = f. \tag{3.6}$$

The coordinates $x_1, \dots, x_k$ of the points of the given stratum S are continuous functions of $y_1, \dots, y_k$, by corollary 3.4.3. This together with (3.6) shows that $\frac{\partial y_l}{\partial y_s}$ are continuous functions of $y_1, \dots, y_k$ on $s^k(\overline{S})$. This proves the proposition. □

3.4.5. Proposition. *We have:*

(1) *The set $(s^k)^{-1}(\partial\Pi^k)$, for $1 \le k \le n$, consists of points of strata of K of dimension at most $k-1$.*
(2) *Every fiber of the projection $\pi : \Pi^k \to \Pi^{k-1}$, ($k \ge 3$), is either a point or a closed interval. If it is an interval, then it contains exactly two points of $\partial\Pi^k$ and these are its endpoints.*
(3) *The k-th coordinates of the upper (or the lower) ends of the fibers of $\pi : \Pi^k \to \Pi^{k-1}$ are continuous functions on Π^{k-1}.*

Proof. To (1): From corollary 3.4.3 follows that a point internal to an l-dimensional stratum of K with $l \ge k$ is mapped by s^k onto a point internal to Π^k. Hence (1).

To (2): If $k \ge 3$, then all fibers of $\pi : \Pi^k \to \Pi^{k-1}$ are compact, which follows from the form of the Newton polynomial s_2. Suppose the fiber over the point $(y_1^0, \dots, y_{k-1}^0)$ is disconnected. Then $\Pi^n \cap \{y_j = y_j^0, 1 \le j \le k-1\}$ is disconnected as well. This contradicts corollary 3.4.3. So each fiber of π is compact and connected, consequently, either a point or a closed interval. Let us assume that it is an interval. It follows from proposition 3.4.4 and (1) that each fiber of π contains only finitely many points belonging to $\partial\Pi^k$. Suppose that a point a, which is internal to the fiber, is a boundary point of Π^k. Choose two points b and c on this fiber, internal to Π^k and such that the open interval bounded by b and c contains a. Consider neighborhoods $U(b)$ and $U(c)$ of b and c in $\mathbb{R}^k$ such that $U(b)$ and $U(c)$ are contained in the interior of Π^k. Let us denote by $U'(b)$ and $U'(c)$ their projection to the first $k-1$ coordinates, and by ω the set of all fibers of π whose projection to the first $k-1$ coordinates lies in $U'(b) \cap U'(c)$. All fibers in ω are intervals intersecting $U(b)$ and $U(c)$. That means that ω contains a neighborhood of a in $\mathbb{R}^k$, i.e., a lies inside Π^k. This proves (2).

To (3): By (2), we may write $\partial\Pi^k = \alpha \cup \beta$, where α is the set of the upper and β the set of the lower endpoints of the fibers of π, then $\alpha \cap \beta$ is the set of the fibers consisting of one point. The boundary $\partial\Pi^k$ is closed. We will show that each of the sets α and β is also closed. Consider a sequence $(a_j)_j \subseteq \alpha$ with $a_j \to b \in \partial\Pi^k$. If $b \notin \alpha$, then b is the lower endpoint of an interval ba, namely, the fiber of π lying over $\pi(b)$. Moreover, $\pi(a_j) \to \pi(b)$. For appropriate points $b', a' \in ba$ satisfying $y_k(b) < y_k(b') < y_k(a') < y_k(a)$, where y_k is the k-th coordinate function, and for an appropriate neighborhood U of $\pi(b)$ in $\mathbb{R}^{k-1}$ the set Π^k will contain the cylinder $b'a' \times U$, which follows from (2). Therefore we would have $y_k(a_j) > y_k(a') > y_k(b)$, a contradiction. This proves that α is closed, for β the proof is analogous. Statement (3) now follows from the fact that a locally bounded function on a locally bounded compact set is continuous if and only if its graph is closed. □

3.4.6. Theorem. *The space $\mathrm{Hyp}_n^0 = \mathrm{Hyp}_n \cap \{a_1 = 0, |a_2| \le 1\}$ is Whitney regular.*

Proof. For any two points $x, y \in \mathrm{Hyp}_n^0$ we construct a piecewise smooth curve $\gamma \subseteq \mathrm{Hyp}_n^0$ joining x and y and consisting of a finite number of algebraic arcs such that $l(\gamma) \le C\|x-y\|$, where $C \ge 1$ does not depend on the choice of x and y.

The construction of γ is carried out in $n-1$ steps. In the k-th step we construct $\gamma^{k+1} \subseteq \Pi_0^{k+1}$, i.e., the projection of γ to the first $k+1$ coordinates, and we show that it satisfies

$$l(\gamma^{k+1}) \leq C_{k+1}\|x^{k+1} - y^{k+1}\|,$$

where $C_{k+1} \geq 1$ does not depend on x^{k+1} and y^{k+1}, the projections to the first $k+1$ coordinates of x and y. This will prove that the set Π_0^{k+1} is Whitney regular.

Let $k=1$. The set Π_0^2 is a closed interval. For γ^2 we may take the straight line connecting x^2 and y^2, and $C_2 = 1$.

Suppose that γ^k has been constructed, for $2 \leq k \leq n-1$. For any point $z \in \gamma^k$ we denote by $L(z)$ the straight line through z which is parallel to the x_{k+1}-axis. Then for every $z \in \gamma^k$ the intersection $\Pi_0^{k+1} \cap L(z)$ is a closed segment with endpoints $a(z)$ and $b(z)$ (or a point if $a(z) = b(z)$). By proposition 3.4.5, the functions ϕ and ψ which assign to $z \in \gamma^k$ the $(k+1)$-st coordinate of the points $a(z)$ and $b(z)$, respectively, are continuous on γ^k.

We claim that ϕ and ψ are Lipschitz functions on γ^k and that their graphs consist of a finite number of algebraic arcs. In fact, the image under s^k of every stratum of K, whose image under s^{k+1} belongs to $\partial\Pi^k$, is a semialgebraic set. Its intersection with any of the algebraic arcs building γ^k is a finite number of algebraic arcs or points. The functions ϕ and ψ are continuously differentiable on the closure of each of the arcs building γ^k, i.e., they are Lipschitz functions on γ^k. It follows from proposition 3.4.4 that the Lipschitz constant d can be chosen not depending on γ^k.

Let us denote by x_{k+1} and y_{k+1} the $(k+1)$-st coordinates of the points x^{k+1} and y^{k+1}. We construct a curve γ_*^{k+1} in Π_0^{k+1} which is the graph of a function on the curve γ^k: Its $(k+1)$-st coordinate equals $\min(\psi(z), \max(x_{k+1}, \phi(z)))$, i.e., the intermediate-value of the numbers $\phi(z), x_{k+1}, \psi(z)$, where $z \in \gamma^k$. It follows from above that γ_*^{k+1} is a Lipschitz function on γ^k, the Lipschitz constant being not greater than d, and that γ_*^{k+1} is piecewise algebraic. Consequently,

$$l(\gamma_*^{k+1}) \leq (1+d^2)^{\frac{1}{2}} l(\gamma^k).$$

One of the ends of γ_*^{k+1} is x^{k+1}. Denote by y' its other end and put $\gamma^{k+1} = \gamma_*^{k+1} \cup y'y^{k+1}$, where $y'y^{k+1}$ is the segment joining y' and y^{k+1}. Then γ^{k+1} is piecewise algebraic and

$$\begin{aligned} l(\gamma^{k+1}) &= l(\gamma_*^{k+1}) + l(y'y^{k+1}) \\ &\leq (1+d^2)^{\frac{1}{2}} l(\gamma^k) + |x_{k+1} - y_{k+1}| \\ &\leq (1+d^2)^{\frac{1}{2}} C_k \|x^k - y^k\| + |x_{k+1} - y_{k+1}| \\ &\leq C_{k+1}\|x^{k+1} - y^{k+1}\|, \end{aligned}$$

where $C_{k+1} := (1+d^2)^{\frac{1}{2}} C_k + 1$ does not depend on x^{k+1} and y^{k+1}. □

Remark. The hypothesis that Hyp_n^0 is Whitney regular has been stated independently by G. Barbançon and by M.D. Bronshtein. Whitney regularity is in connection with the Whitney extension theorem [**Whi34**]. In fact, see [**Bal84**] and [**Bar72**], we have: *Every symmetric function $f : \mathbb{R}^n \to \mathbb{R}$ of class C^{nr} can be expressed as a C^r function of $s_i = \sum_{j=1}^n x_j^i$ $(1 \leq i \leq n)$ or, equivalently, of $\sigma_i = \sum_{1 \leq j_1 < \cdots < j_i \leq n} x_{j_1} \cdots x_{j_i}$ $(1 \leq i \leq n)$. The estimation of the smoothness is exact.*

Remark. There is a more general notion of regularity ([**Pfl01**], [**Tou72**]): A compact subset $K \in \mathbb{R}^n$ is called *p-regular*, for real $p \geq 1$, if it is finitely arc-wise connected, and there exists a constant $C > 0$ such that for all $x, y \in K$

$$\delta(x, y) \leq C\|x - y\|^{\frac{1}{p}},$$

where δ denotes the *geodesic distance*

$$\delta(x, y) = \inf\{l(\gamma) : \gamma \in C([0, 1], K), \gamma(0) = x, \gamma(1) = y\}.$$

A locally closed, connected, and locally finitely arc-wise connected set $A \subseteq \mathbb{R}^n$ is called *p-regular*, if any point in A has a compact p-regular neighborhood. If for any point $z \in A$ there exists a $p \geq 1$, depending on z, and a p-regular compact neighborhood $K \subseteq \mathbb{R}^n$, then we say A is *Whitney-Tougeron regular*.

There is the following result: *Let X be a compact connected subanalytic subset of $\mathbb{R}^n$. Then there is a positive integer p such that X is p-regular (where the curves can be chosen semianalytic).* See [**Sta82**]. Closed *subanalytic* sets are precisely the images of real analytic sets by proper real analytic mappings.

Note that theorem 3.4.6 is a special case of following result due to Barbançon [**Bar05**]: *Let $G \in \mathrm{O}(V)$ be a finite reflection group and $\sigma = (\sigma_1, \ldots, \sigma_n) : V \to \mathbb{R}^n$, where $\sigma_1, \ldots, \sigma_n$ are homogeneous generators of the algebra $\mathbb{R}[V]^G$ of invariant polynomials. Then any compact subset of the orbit space $V/G \cong \sigma(V) \subseteq \mathbb{R}^n$ is 1-regular.* Note that 1-regularity of $\sigma(V)$ does not depend on the choice of the generators $\sigma_1, \ldots, \sigma_n$, since any two choices differ by a polynomial diffeomorphism.

CHAPTER 4

The approach of Alekseevsky, Kriegl, Losik, and Michor

The present chapter is devoted to a well structured approach to the problem of choosing roots of hyperbolic polynomials smoothly. It is due to Alekseevsky, Kriegl, Losik, and Michor [**AKLM98**]. The last section 4.6 gives a short glance to the complex case, where there are no restrictions on the roots to be real.

4.1. Choosing differentiable square roots

As introduction let us investigate the case of quadratic hyperbolic polynomials $P(t)(x) = x^2 - a_1(t)x + a_2(t)$ depending on a real parameter t. By replacing the variable x with $y = x - \frac{a_1(t)}{2}$, we reduce the problem to $a_1 = 0$.

4.1.1. Proposition. *Consider $P(t)(x) = x^2 - f(t)$ for a non-negative function f defined on an open interval.*

If f is smooth and it is nowhere flat of infinite order (see definition 4.2.4), then smooth roots of P exist.

If f is of class C^2, then C^1 roots exist.

If f is of class C^4, then twice differentiable roots exist.

Proof. Suppose f is smooth and nowhere flat of infinite order, and consider an arbitrary point t_0 in the domain of definition of f. If $f(t_0) > 0$, then we have obvious local smooth roots $\pm\sqrt{f(t)}$. If $f(t_0) = 0$, we have to find a smooth function x such that $f = x^2$, a smooth square root of f. Since f is not flat at t_0 and always non-negative, the first non-zero derivative at t_0 has even order $2m$ and is positive. We have $f(t) = (t-t_0)^{2m} f_{2m}(t)$, where $f_{2m}(t) := \int_0^1 \frac{(1-r)^{2m-1}}{(2m-1)!} f^{(2m)}(t_0+r(t-t_0))dr$ by means of Taylor's formula. Now, f_{2m} is a smooth function with $f_{2m}(t_0) = \frac{1}{(2m)!} f^{(2m)}(t_0) > 0$. Then, $x(t) := (t-t_0)^m \sqrt{f_{2m}(t)}$ is a local smooth root. Since t_0 was arbitrary, we have found local smooth roots everywhere. One can piece them together in order to get global smooth roots, changing sign at all points, where the first non-vanishing derivative of f is of order $2m$ with m odd. These points are discrete. This shows the first assertion in the proposition.

Let us consider now a non-negative function f of class C^2. We claim that the equation $x^2 = f(t)$ admits a C^1 solution $x(t)$, globally in t. Let t_0 be fixed. If $f(t_0) > 0$, then there is locally even a C^2 solution $x_\pm(t) = \pm\sqrt{f(t)}$. If $f(t_0) = 0$, then, f being non-negative, we have $f(t) = (t-t_0)^2 h(t)$, where $h \ge 0$ is continuous everywhere and C^2 off t_0 with $h(t_0) = \frac{1}{2} f''(t_0)$. For $h(t_0) > 0$, put $x_\pm(t) = \pm(t-t_0)\sqrt{h(t)}$ which is C^2 off t_0, and

$$x'_\pm(t_0) = \lim_{t\to t_0} \frac{x_\pm(t) - x_\pm(t_0)}{t - t_0} = \lim_{t\to t_0} \pm\sqrt{h(t)} = \pm\sqrt{h(t_0)} = \pm\sqrt{\frac{1}{2} f''(t_0)}.$$

For $h(t_0) = 0$, we choose $x_\pm(t_0) = 0$, and any choice of the roots is then differentiable at t_0 with derivative 0, by the same calculation.

One can piece together these local roots: At zeros t of f where $f''(t) > 0$ our root has to pass through 0 (examine $x'_\pm$), but, for t where $f''(t) = 0$, the choice of the root does not matter. The set $\{t : f(t) = f''(t) = 0\}$ is closed, so its complement is a disjoint union of open intervals. Choose a point in each of these intervals, where $f(t) > 0$, and start there with the positive root x_+, changing signs at points, where $f(t) = 0 \neq f''(t)$; these points do not accumulate in the intervals. Hence, we get a differentiable choice of a root $x(t)$ on each of this open intervals which extends to a global differentiable root which is 0 on $\{t : f(t) = f''(t) = 0\}$, by the observation at the beginning of this paragraph.

Note that for this global differentiable root x we have

$$x'(t) = \begin{cases} \frac{f'(t)}{2x(t)} & \text{if } f(t) > 0 \\ \pm\sqrt{\frac{1}{2}f''(t)} & \text{if } f(t) = 0. \end{cases}$$

We have seen that in points t_0 with $f(t_0) > 0$ the root x is C^2. Locally around points t_0 with $f(t_0) = 0$ and $f''(t_0) > 0$ the root x is C^1, since it is even C^2 off t_0 and for $t \neq t_0$ near t_0 we have $f(t) > 0$ and $f'(t) \neq 0$, so by de l'Hospital we get

$$\lim_{t\to t_0} x'(t)^2 = \lim_{t\to t_0} \frac{f'(t)^2}{4f(t)} = \lim_{t\to t_0} \frac{2f'(t)f''(t)}{4f'(t)} = \frac{1}{2}f''(t_0) = x'(t_0)^2,$$

and since the choice of signs was coherent, x' is continuous at t_0. Finally, if $f(t_0) = 0$ and $f''(t_0) = 0$, then $x'(t_0) = 0$, and $x'(t) \to 0$ for $t \to t_0$ for both expressions of x' given above, by lemma 4.1.2 below. Thus, x is of class C^1.

To prove the third part of the proposition, where $f \geq 0$ is C^4, we modify the C^1 root from above to be twice differentiable. Near points t_0 with $f(t_0) > 0$ any continuous root $x_\pm = \pm\sqrt{f(t)}$ is even C^4. Near points t_0 with $f(t_0) = f'(t_0) = 0$ we have $f(t) = (t-t_0)^2h(t)$, where $h(t) := \int_0^1 (1-r)f''(t_0+r(t-t_0))dr$ is non-negative and C^2. It follows that $h''(t_0) = \frac{1}{12}f^{(4)}(t_0)$. We may choose a C^1 solution z of the equation $z^2 = h$ by the arguments above, then $z'(t_0) = \pm\sqrt{\frac{1}{2}h''(t_0)}$. Consequently, $x(t) := (t - t_0)z(t)$ is twice differentiable at t_0, since

$$\frac{x'(t) - x'(t_0)}{t - t_0} = \frac{z(t) + (t - t_0)z'(t) - z(t_0)}{t - t_0} = z'(t) + \frac{z(t) - z(t_0)}{t - t_0}$$

which converges to

$$2z'(t_0) = \pm 2\sqrt{\frac{1}{2}h''(t_0)} = \pm 2\sqrt{\frac{1}{4!}f^{(4)}(t_0)},$$

as $t \to t_0$. If $f(t_0) = f''(t_0) = f^{(4)}(t_0) = 0$, then any C^1 choice of the roots is twice differentiable at t_0, by the previous calculation, in particular $x(t) = |t - t_0|z(t)$.

Let us piece together these solutions similarly as above. Suppose y is a global C^1 root of $x^2 = f$, chosen as before changing sign only at points t with $f(t) = 0 < f''(t)$. It is easy to see that this choice provides a twice differentiable root of f. □

4.1.2. LEMMA. *Let f be a non-negative C^2 function with $f(t_0) = 0$ for a point t_0 in $\mathbb{R}$. Then, for all $t \in \mathbb{R}$, we have*

$$f'(t)^2 \leq 2f(t)\max\{f''(t_0 + r(t - t_0)) : 0 \leq r \leq 2\}. \tag{4.1}$$

PROOF. Since f is non-negative, $f(t) = 0$ implies $f'(t) = 0$, so (4.1) holds at zeros of f. Hence we assume $f(t) > 0$. We use Taylor's formula

$$f(t + s) = f(t) + f'(t)s + \int_0^1 (1 - r)f''(t + rs)drs^2. \tag{4.2}$$

In particular we get (replacing t by t_0 and then t_0+s by t)

$$\begin{aligned} f(t) &= 0+0+\int_0^1 (1-r)f''(t_0+r(t-t_0))dr(t-t_0)^2 \\ &\le \frac{(t-t_0)^2}{2}\max\{f''(t_0+r(t-t_0)) : 0\le r\le 2\}. \end{aligned} \tag{4.3}$$

Now in (4.2) we replace s by $-\epsilon s$, where $\epsilon = \operatorname{sgn}(f'(t))$ and obtain

$$0 \le f(t-\epsilon s) = f(t) - |f'(t)|s + \int_0^1 (1-r)f''(t-\epsilon r s)drs^2. \tag{4.4}$$

Let us assume $t \ge t_0$ and put

$$s(t) := \sqrt{\frac{2f(t)}{\max\{f''(t_0+r(t-t_0)) : 0\le r\le 2\}}},$$

then, by (4.3) and since $f(t) > 0$ (which implies $\max\{f''(t_0+r(t-t_0)) : 0 \le r \le 2\} > 0$), we have $0 < s(t) \le t-t_0$, and $s(t)$ is well defined. This choice of s in (4.4) gives

$$\begin{aligned} |f'(t)| &\le \frac{1}{s(t)}\left(f(t) + s(t)^2\int_0^1 (1-r)f''(t-\epsilon r s(t))dr\right) \\ &\le \frac{1}{s(t)}\left(f(t) + \frac{s(t)^2}{2}\max\{f''(t-\epsilon r s(t)) : 0\le r\le 1\}\right) \\ &\le \frac{1}{s(t)}\left(f(t) + \frac{s(t)^2}{2}\underbrace{\max\{f''(t-r(t-t_0)) : -1\le r\le 1\}}_{=\max\{f''(t_0+r(t-t_0)):0\le r\le 2\}}\right) \\ &= \frac{2f(t)}{s(t)} = \sqrt{2f(t)\max\{f''(t_0+r(t-t_0)) : 0\le r\le 2\}} \end{aligned}$$

which proves (4.1) for $t \ge t_0$. Since the assertion is symmetric, it then holds for all $t \in \mathbb{R}$. □

4.1.3. Remark. The second statement in proposition 4.1.1 has been already found by Mandai [**Man85**]. Lemma 4.1.2 can be found in a similar version in [**Die70**].

4.1.4. Note that the differentiability assumptions imposed on f in proposition 4.1.1 are best possible: If they are slightly weakened, then the statements are false.

Examples. If $f \ge 0$ is only C^1, then there may not exist a differentiable root of $x^2 = f(t)$. For instance, the function $f(t) := t^2\sin^2(\log t)$ is C^1, but the square roots $\pm t\sin(\log t)$ are not differentiable at 0.

If $f \ge 0$ is twice differentiable, there may not exist a C^1 root. E.g., $f(t) = t^4\sin^2(\frac{1}{t})$ is twice differentiable, but $\pm t^2\sin(\frac{1}{t})$ is differentiable but not C^1.

If $f \ge 0$ is only C^3, then there may not exist a twice differentiable root of $x^2 = f(t)$. E.g., $f(t) = t^4\sin^2(\log t)$ is C^3, but $\pm t^2\sin(\log t)$ is only C^1 and not twice differentiable.

If $f \ge 0$ is smooth but flat at 0, in general the equation $x^2 = f(t)$ has no C^2 solution as the following example shows, which is an application of the general curve lemma ([**KM97**, Chapter III, 12.2]): Let $h : \mathbb{R} \to [0,1]$ be smooth with $h(t) = 1$ for $t \ge 0$ and $h(t) = 0$ for $t \le -1$. Then, we claim that the function

$$f(t) := \sum_{n=1}^{\infty} h_n(t-t_n)\cdot\left(\frac{2n}{2^n}(t-t_n)^2 + \frac{1}{4^n}\right), \tag{4.5}$$

where

$$h_n(t) := h\left(n^2\left(\frac{1}{n\cdot 2^{n+1}}+t\right)\right)\cdot h\left(n^2\left(\frac{1}{n\cdot 2^{n+1}}-t\right)\right)$$

and

$$t_n := \sum_{k=1}^{n-1}\left(\frac{2}{k^2}+\frac{2}{k\cdot 2^{k+1}}\right)+\frac{1}{n^2}+\frac{1}{n\cdot 2^{n+1}},$$

is non-negative and is smooth. It is a direct consequence of the fact that the sum on the right-hand side of (4.5) consists of at most one summand for each t, and that the derivatives of the summands converge uniformly to 0. This in turn is seen as follows: Observe that $h_n(t) = 1$ for $|t| \le \frac{1}{n\cdot 2^{n+1}}$ and $h_n(t) = 0$ for $|t| \ge \frac{1}{n\cdot 2^{n+1}} + \frac{1}{n^2}$, hence $h_n(t-t_n) \ne 0$ only for $r_n < t < r_{n+1}$, where $r_n := \sum_{k=1}^{n-1}\left(\frac{2}{k^2}+\frac{2}{k\cdot 2^{k+1}}\right)$, which shows the first statement. To prove the second statement let $c_n(s) := \frac{2n}{2^n}s^2 + \frac{1}{4^n} \ge 0$ and $H_i := \sup\{|h^{(i)}(t)| : t \in \mathbb{R}\}$. Then,

$$\begin{aligned} n^2\sup\{|(h_n\cdot c_n)^{(k)}(t)| : t\in\mathbb{R}\} &= n^2\sup\left\{|(h_n\cdot c_n)^{(k)}(t)| : |t| \le \frac{1}{n\cdot 2^{n+1}}+\frac{1}{n^2}\right\} \\ &\le n^2\sum_{i=0}^{k}\binom{k}{i}n^{2i}H_i\sup\left\{|c_n^{(k-i)}(t)| : |t|\le \frac{1}{n\cdot 2^{n+1}}+\frac{1}{n^2}\right\} \\ &\le \left(\sum_{i=0}^{k}\binom{k}{i}n^{2i+2}H_i\right)\sup\{|c_n^{(j)}(t)| : |t|\le 2, j\le k\}, \end{aligned} \tag{4.6}$$

since we have $|h_n^{(i)}(t)| \le n^{2i}H_i$ for all t. Note that c_n is rapidly decreasing in $C^\infty(\mathbb{R},\mathbb{R})$, i.e., $\{p(n)c_n : n \in \mathbb{N}\}$ is bounded in $C^\infty(\mathbb{R},\mathbb{R})$ for each polynomial p, therefore, the right-hand side of inequality (4.6) is bounded with respect to $n \in \mathbb{N}$. Consequently, the series $\sum_n h_n(\ \ -t_n)c_n(\ \ -t_n)$ converges uniformly in each derivative, and thus represents an element of $C^\infty(\mathbb{R},\mathbb{R})$.

Moreover, we have

$$f(t_n) = \frac{1}{4^n},\ f'(t_n) = 0, \quad \text{and} \quad f''(t_n) = \frac{2n}{2^{n-1}}.$$

Let us assume that $f(t) = g(t)^2$ for t near $\sup_n t_n < \infty$, where g is twice differentiable. Then

$$\begin{aligned} f' &= 2gg' \\ f'' &= 2gg'' + 2(g')^2 \\ 2ff'' &= 4g^3g'' + (f')^2 \\ 2f(t_n)f''(t_n) &= 4g(t_n)^3g''(t_n) + f'(t_n)^2 \\ 2\cdot\frac{1}{4^n}\frac{2n}{2^{n-1}} &= \pm 4\left(\frac{1}{4^n}\right)^{\frac{3}{2}}g''(t_n), \end{aligned}$$

whence $g''(t_n) = \pm 2n$. So g cannot be C^2, and g' cannot satisfy a local Lipschitz condition near $\lim t_n$.

4.1.5. Remarks. (1): Note that there are results concerning higher-dimensional parameter spaces: In [**Gla63b**] Glaeser proved that a non-negative C^2 function on an open subset of $\mathbb{R}^n$ which vanishes of second order has a positive square root of class C^1. Moreover, in that paper a smooth non-negative function $f : \mathbb{R} \to \mathbb{R}$ is constructed which is flat at 0 such that the positive square root is not C^2. In [**Die70**] Dieudonné gave shorter proofs of Glaeser's results.

(2): A function g of a certain regularity satisfying $f = g^2$ is often referred to as an *admissible square root* of f. There is a recent work by Bony, Broglia, Colombini, and Pernazza [**BBCP06**] partly dedicated to the study of the regularity of the admissible square root of a non-negative regular function in one dimension. We indicate shortly the main results:

- Given any modulus of continuity ω there are non-negative smooth functions such that the first derivative of any of their admissible square roots is not ω-continuous.
- A non-negative C^4 function which takes the value 0 at all its local minima has an admissible square root in $C^2(\mathbb{R})$.
- A non-negative C^4 function f has an admissible C^2 square root if and only if there exists a continuous function γ vanishing on the set of points where f is flat, i.e., $\{t \in \mathbb{R} : f^{(k)}(t) = 0 \text{ for all } 0 \le k \le 4\}$, such that, for any local minimum t_0 of f where $f(t_0) > 0$, we have $f''(t_0) \le \gamma(t_0) f(t_0)^{\frac{1}{2}}$.
- For any given modulus of continuity ω there is a non-negative smooth function on $\mathbb{R}$, taking the value 0 at all its local minima, which has no admissible square root of class $C^{2,\omega}$.

DEFINITION. A *modulus of continuity* is a continuous increasing concave function ω, defined on an interval $[0, t_0]$ and satisfying $\omega(0) = 0$. If Ω is an open subset of $\mathbb{R}^n$, a function $f : \Omega \to \mathbb{R}$ will be called *ω-continuous* on Ω if the quantity

$$[f]_\omega = \sup \frac{|f(y) - f(x)|}{\omega(|y - x|)}$$

is finite, where the supremum is over all $0 < |x - y| < \min\{t_0, \frac{1}{2} \operatorname{dist}(x, \mathbb{R}^n \backslash \Omega)\}$. For a non-negative integer k we say that f is of class $C^{k,\omega}$ if it belongs to C^k and if

$$\|f\|_{k,\omega} = \|f\|_{C^k} + \sum_{|\alpha|=k} \left[\frac{\partial^\alpha f}{\partial x^\alpha}\right]_\omega$$

is finite. For every continuous function f there exists a modulus of continuity ω such that f is ω-continuous.

4.2. Factorizing the curve of polynomials

In this section we present a well structured approach to the problem of choosing roots of polynomials smoothly. At its end we shall dispose of a effective algorithm which yields a factorization of a curve of hyperbolic polynomials in solvable and potentially unsolvable part. That means that the latter part of the factorization may or may not be solvable in the sense introduced in the following definition. We shall give an example at the end of this section. We use the notation of section 3.1.

4.2.1. Regular solvability. Let us consider a smooth curve of hyperbolic polynomials

$$P(t)(x) = x^n - a_1(t)x^{n-1} + \cdots + (-1)^n a_n(t).$$

DEFINITION. We will say that the smooth curve of polynomials P is *smoothly solvable near* t_0, if there exist n smooth functions $x_1(t), \ldots, x_n(t)$ of the parameter t defined near $t_0 \in \mathbb{R}$ which parameterize the roots of $P(t)$ for each t. In analogy we shall use to say that P is *real analytically solvable near* t_0 (or *holomorphically solvable near* t_0), if there are real analytic (or holomorphic) functions $x_1(t), \ldots, x_n(t)$ defined near $t_0 \in \mathbb{R}$ (or $t_0 \in \mathbb{C}$) which parameterize the roots of $P(t)$ for each t. (In the latter case there is no restriction on the roots of P, i.e., the polynomials $P(t)$ need not be hyperbolic.)

Note that the problem of smooth solvability of P can be reduced to the case $a_1 = 0$, replacing the variable x with the variable $y = x - \frac{a_1(t)}{n}$. We shall use this reduction in the following whenever it is meaningful and yields a simplification.

4.2.2. All roots distinct. First we treat the case when all roots of $P(t_0)$ are distinct. Without loss of generality we may assume that $t_0 = 0$.

PROPOSITION. *Let P be a smooth curve of hyperbolic polynomials as above such that the roots of $P(0)$ are all distinct. Then P is smoothly solvable near 0.*

This is also true in the real analytic case and for higher dimensional parameters, and in the holomorphic case for complex roots.

PROOF. Let $x_1, \dots, x_n$ be the roots of $P(0)$ and write $P(0)(x) = \prod_{i=1}^n (x - x_i)$. The derivative $\frac{d}{dx} P(0)(x) = \sum_{i=1}^n (x - x_1) \cdots \widehat{(x - x_i)} \cdots (x - x_n)$ does not vanish at any root $x_1, \dots, x_n$, since they are distinct. Thus, by the implicit function theorem, we have local smooth solutions $x_1(t), \dots, x_n(t)$ of $P(t, x) = P(t)(x) = 0$ with $x_1(0) = x_1, \dots, x_n(0) = x_n$.

The same arguments work in the cases listed in the second part of the proposition. $\square$

4.2.3. LEMMA (Splitting lemma). *Let $P_0 = x^n - a_1 x^{n-1} + \dots + (-1)^n a_n$ be a polynomial satisfying $P_0 = P_1 \cdot P_2$, where P_1 and P_2 are polynomials without common root. Then for P near P_0 we have $P = P_1(P) \cdot P_2(P)$ for real analytic mappings of monic polynomials $P \mapsto P_1(P)$ and $P \mapsto P_2(P)$, defined for P near P_0, with the given initial values.*

PROOF. Let the polynomial P_0 be represented as the product

$$P_0 = P_1 \cdot P_2 = (x^p - b_1 x^{p-1} + \dots + (-1)^p b_p) \cdot (x^q - c_1 x^{q-1} + \dots + (-1)^q c_q),$$

where $p + q = n$. Let $x_1, \dots, x_n$ be the roots of P_0, ordered in such a way that the first p are the roots of P_1 and the last q are those of P_2. Then $(a_1, \dots, a_n) = \Phi^{p,q}(b_1, \dots, b_p, c_1, \dots, c_q)$ for a polynomial mapping $\Phi^{p,q}$ and we get

$$\sigma^n = \Phi^{p,q} \circ (\sigma^p \times \sigma^q)$$

and

$$\det(d\sigma^n) = \det(d\Phi^{p,q}(b, c)) \det(d\sigma^p) \det(d\sigma^q),$$

where $b = (b_1, \dots, b_p)$ and $c = (c_1, \dots, c_q)$ and $\sigma^n, \sigma^p, \sigma^q$ are as in 3.1.1. From section 3.1 we conclude

$$\prod_{1 \le i < j \le n} (x_i - x_j) = \det(d\Phi^{p,q}(b, c)) \prod_{1 \le i < j \le p} (x_i - x_j) \prod_{p+1 \le i < j \le n} (x_i - x_j)$$

which in turn implies

$$\det(d\Phi^{p,q}(b, c)) = \prod_{1 \le i \le p < j \le n} (x_i - x_j) \neq 0,$$

since P_1 and P_2 do not have common roots. So, by the inverse function theorem, $\Phi^{p,q}$ is a real analytic diffeomorphism near (b, c). $\square$

4.2.4. Multiplicity.

DEFINITION. For a continuous function f defined near 0 in $\mathbb{R}$ let the *multiplicity* or *order of flatness* $m(f)$ *at* 0 be the supremum of all integers p such that $f(t) = t^p g(t)$ near 0 for a continuous function g.

Similarly one can define the multiplicity of a function at any $t \in \mathbb{R}$. Note that, if f is of class C^n and $m(f) < n$, then $f(t) = t^{m(f)} g(t)$ near 0, where now g is $C^{n-m(f)}$ and $g(0) \neq 0$.

If f is a continuous function on the space of polynomials, then for a fixed continuous curve P of polynomials we will denote by $m(f)$ the multiplicity at 0 of $t \mapsto f(P(t))$.

4.2.5. PROPOSITION. *Suppose that the smooth curve of polynomials*

$$P(t)(x) = x^n + a_2(t)x^{n-2} - \cdots + (-1)^n a_n(t)$$

is smoothly solvable with smooth roots $t \mapsto x_i(t)$*, where* $1 \leq i \leq n$*, and that all roots of* $P(0)$ *are equal. Then, for all* $2 \leq k \leq n$ *we have*

$$m(\tilde{\Delta}_k) \geq k(k-1) \min_{1\leq i\leq n} m(x_i)$$

and

$$m(a_k) \geq k \min_{1\leq i\leq n} m(x_i).$$

Note that the $\tilde{\Delta}_k$ *are defined in 3.1.1. This result holds in the real analytic case and in the holomorphic case, too.*

PROOF. The second inequality stated in the proposition follows from

$$a_k(t) = \sigma_k(x_1(t), \dots, x_n(t)) = \sum_{1\leq j_1<\cdots<j_k\leq n} x_{j_1}(t)\cdots x_{j_k}(t).$$

Observe that in equation (3.3) each summand on the right-hand side has exactly $k(k-1)$ linear factors in the x_i, hence we get the other inequality. The real analytic case and the holomorphic case can be treated in the same way, because the two equations used in the proof remain valid. □

4.2.6. LEMMA. *Let* $P(x) = x^n - a_1x^{n-1} + \cdots + (-1)^n a_n$ *be a hyperbolic polynomial of degree* n*. If* $a_1 = a_2 = 0$*, then all roots of* P *are equal to zero.*

PROOF. From (3.1) we have $\sum_{j=1}^n x_j^2 = s_2(x) = \sigma_1^2(x) - 2\sigma_2(x) = a_1^2 - 2a_2 = 0$, where $x_1, \dots, x_n$ are the roots of P and $x = (x_1, \dots, x_n)$. Since the roots are real, the lemma follows. □

Note that the assumption on the roots of P of being real is the crucial point in the proof. The lemma does not hold when no restrictions are made on the roots.

4.2.7. LEMMA (Multiplicity lemma). *Consider a smooth curve of hyperbolic polynomials*

$$P(t)(x) = x^n + a_2(t)x^{n-2} - \cdots + (-1)^n a_n(t).$$

Then, for each integer r*, the following conditions are equivalent:*

(1) $m(a_k) \geq kr$*, for all* $2 \leq k \leq n$*;*
(2) $m(\tilde{\Delta}_k) \geq k(k-1)r$*, for all* $2 \leq k \leq n$*;*
(3) $m(a_2) \geq 2r$*.*

PROOF. We only have to treat $r > 0$.

(1) ⇒ (2): From (3.1) we deduce (by induction) that $m(\tilde{s}_k) \geq kr$ for all $k \geq 0$, where $\tilde{s}_k$ is defined us usual by $s_k = \tilde{s}_k \circ \sigma^n$. Hence, observing that

$$\tilde{\Delta}_k = \det(\tilde{B}_k) = \det \begin{pmatrix} \tilde{s}_0 & \tilde{s}_1 & \dots & \tilde{s}_{k-1} \\ \tilde{s}_1 & \tilde{s}_2 & \dots & \tilde{s}_k \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{s}_{k-1} & \tilde{s}_k & \dots & \tilde{s}_{2k-2} \end{pmatrix}$$

is a polynomial in variables $\tilde{s}_i$, where in each summand the indices add up to $k(k-1)$, we obtain (2).

(2) ⇒ (3): This is clear, since

$$\tilde{\Delta}_2 = \det \begin{pmatrix} \tilde{s}_0 & \tilde{s}_1 \\ \tilde{s}_1 & \tilde{s}_2 \end{pmatrix} = \det \begin{pmatrix} n & a_1 \\ a_1 & a_1^2 - 2a_2 \end{pmatrix} = \det \begin{pmatrix} n & 0 \\ 0 & -2a_2 \end{pmatrix} = -2na_2.$$

(3) ⇒ (1): From $a_2(0) = 0$ and lemma 4.2.6 it follows that all roots of the polynomial $P(0)$ are equal to zero and, consequently, $a_3(0) = \cdots = a_n(0) = 0$, too.

Under these conditions near 0 we have $a_2(t) = t^{2r} a_{2,2r}(t)$ and $a_k(t) = t^{m_k} a_{k,m_k}(t)$ for $3 \le k \le n$, where the m_k are positive integers and $a_{2,2r}, a_{3,m_3}, \dots, a_{n,m_n}$ are smooth functions, and where we may assume that either $m_k = m(a_k) < \infty$ or, if $m(a_k) = \infty$, that $m_k \ge kr$.

Let us suppose indirectly that for some $k > 2$ we have $m_k = m(a_k) < kr$. We put

$$m := \min\left(r, \frac{m_3}{3}, \dots, \frac{m_n}{n}\right) < r. \tag{4.7}$$

We consider the following continuous curve of polynomials for (small) $t \ge 0$:

$$\begin{aligned}\bar{P}_m(t)(x) := x^n &+ a_{2,2r}(t) t^{2r-2m} x^{n-2} \\ &- a_{3,m_3}(t) t^{m_3 - 3m} x^{n-3} + \dots + (-1)^n a_{n,m_n}(t) t^{m_n - nm}.\end{aligned} \tag{4.8}$$

It is easy to see that $\bar{P}_m(t)(x) = t^{-nm} P(t)(t^m x)$, for $t > 0$. So, if $x_1(t), \dots, x_n(t)$ are the real roots of $P(t)$, then $t^{-m} x_1(t), \dots, t^{-m} x_n(t)$ are those of $\bar{P}_m(t)$, for $t > 0$. Consequently, $\{\bar{P}_m(t) : t > 0\}$ is a family of hyperbolic polynomials. Since by theorem 3.1.2 the space of hyperbolic polynomials of a fixed degree is closed, $\bar{P}_m(0)$ is also a polynomial with all roots real.

By lemma 4.2.6, all roots of the polynomial $\bar{P}_m(0)$ are equal to zero, and for those k with $m_k = km$ we find $a_{k,m_k}(0) = 0$. Therefore, $m(a_k) > m_k$ for those k, a contradiction. □

The essence of the multiplicity lemma remains true, if the differentiability assumptions on the curve of polynomials are weakened. Since we shall need this stronger form of the multiplicity lemma later on, we want to discuss it here in detail.

4.2.8. Lemma (Strong multiplicity lemma). *Consider a curve of hyperbolic polynomials*

$$P(t)(x) = x^n + a_2(t) x^{n-2} - \dots + (-1)^n a_n(t),$$

where a_k is of class C^k for all $2 \le k \le n$. Then the following two conditions are equivalent:

(1) *$a_k(t) = t^k a_{k,k}(t)$ near 0 for a continuous function $a_{k,k}$, for $2 \le k \le n$;*
(2) *$a_2(t) = t^2 a_{2,2}(t)$ near 0 for a continuous function $a_{2,2}$.*

Proof. To show the nontrivial implication (2) $\Rightarrow$ (1) we simply follow the third part of the foregoing proof with $r = 1$ and change it slightly. By lemma 4.2.6 we find that all coefficients of P vanish at $t = 0$. So, near 0 we have $a_2(t) = t^2 a_{2,2}(t)$ and $a_k(t) = t^{m_k} a_{k,k}(t)$ for $3 \le k \le n$, where we define $m_k := \min(k, m(a_k))$ for all k. Then the m_k are positive integers such that $m_k \le k$. And the functions $a_{3,3}, \dots, a_{n,n}$ are continuous, because $a_k \in C^k$ for $3 \le k \le n$.

Now suppose for contradiction that for some $k > 2$ we have $m_k < k$. We define $m < 1$ as in (4.7) (with $r = 1$) and the continuous curve of polynomials $\bar{P}_m$ as in (4.8). By the same arguments we find that all roots of $\bar{P}_m(0)$ vanish, and hence for those k with $m_k = km$ we have $a_{k,k}(0) = 0$. But then $m_k = m(a_k)$ for these k, a contradiction. □

4.2.9. Factorizing algorithm. The preparatory work we have done so far allows us now to present the announced algorithm for a factorization of a curve of hyperbolic polynomials in solvable and potentially unsolvable part.

Algorithm. Consider a smooth curve of hyperbolic polynomials

$$P(t)(x) = x^n - a_1(t) x^{n-1} + a_2(t) x^{n-2} - \dots + (-1)^n a_n(t).$$

The algorithm consists of the following steps:

(1) If all roots of $P(0)$ are pairwise different, P is smoothly solvable for t near 0, by proposition 4.2.2.

(2) If there are distinct roots at $t = 0$, we put them into two disjoint subsets which splits $P(t) = P_1(t) \cdot P_2(t)$ near 0 by the splitting lemma 4.2.3. We then feed $P_1(t)$ and $P_2(t)$ (which have lower degrees) into the algorithm.

(3) If all roots of $P(0)$ are equal, then we first reduce $P(t)$ to the case $a_1 = 0$ by replacing the variable x with $y = x - \frac{a_1(t)}{n}$. Then, by Vieta's formula for a_1, all roots of $P(0)$ are equal to 0. Consequently, a_2 vanishes at 0, i.e., $m(a_2) > 0$.

(3a) If $m(a_2)$ is finite, then it has to be even, since by theorem 3.1.2 the hyperbolicity of P forces a_2 to be non-positive everywhere: $0 \leq \tilde{\Delta}_2 = -2na_2$. We have $m(a_2) := 2r$ for a positive integer r, and from the multiplicity lemma 4.2.7 we obtain $a_k(t) = t^{kr} a_{k,kr}(t)$ near 0 for smooth $a_{k,kr}$ and $2 \leq k \leq n$. Let us consider the following smooth curve of polynomials

$$P_r(t)(x) := x^n + a_{2,2r}(t)x^{n-2} - a_{3,3r}(t)x^{n-3} + \cdots + (-1)^n a_{n,nr}(t).$$

Since $P_r(t)(x) = t^{-nr} P(t)(t^r x)$, the $P_r(t)$ are again hyperbolic polynomials, and, if $P_r(t)$ is smoothly solvable and $x_j(t)$ are its smooth roots, then $t^r x_j(t)$ are the roots of $P(t)$ and hence the original curve P is smoothly solvable, too. Because of $a_{2,2r}(0) \neq 0$, not all roots of $P_r(0)$ are equal (by Vieta's formulas), and we may feed P_r into step (2) of the algorithm.

(3b) If $m(a_2)$ is infinite and $a_2 = 0$, then all roots of P are identically 0 by lemma 4.2.6, and thus P is smoothly solvable.

(3c) Finally, if $m(a_2)$ is infinite and $a_2 \neq 0$, then by the multiplicity lemma 4.2.7 all $m(a_k)$ for $2 \leq k \leq n$ are infinite. In this case we keep $P(t)$ as factor of the original curve of polynomials with all coefficients infinitely flat at $t = 0$, after forcing $a_1 = 0$. This means that all roots of $P(t)$ meet of infinite order of flatness (see definition 4.2.4) at $t = 0$ for any choice of the roots. This can be seen as follows: If $x(t)$ is any root of $P(t)$, then $y(t) = t^{-r} x(t)$ is a root of $P_r(t)$, hence bounded by lemma 2.3.2, so $x(t) = t^{r-1} \cdot ty(t)$, and $t \mapsto ty(t)$ is continuous at $t = 0$.

Evidently this algorithm always stops, since every passing through either yields the desired factorization or lowers the degree of the involved polynomial. It produces a splitting of the original polynomial

$$P(t) = P^{(\infty)}(t) \cdot P^{(s)}(t),$$

where $P^{(\infty)}(t)$ has the property that each root meets another one of infinite order at $t = 0$, and where $P^{(s)}(t)$ is smoothly solvable, and no two of its roots meet of infinite order at $t = 0$, if they are not equal. Any two choices of smooth roots of $P^{(s)}$ differ by a permutation.

By means of an example we demonstrate now that the factor $P^{(\infty)}$ may or may not be smoothly solvable.

Example. For a non-negative smooth function f which is flat at 0 consider the following smooth curve of hyperbolic polynomials

$$P(t)(x) = x^4 - (f(t) + t^2)x^2 + t^2 f(t).$$

Here the algorithm produces the factorization

$$P(t)(x) = (x^2 - f(t)) \cdot (x - t)(x + t)).$$

If, for instance, f has the form $f(t) = g(t)^2$ for a smooth function g, then $P^{(\infty)}(t)(x) = x^2 - f(t)$ is smoothly solvable near 0. For the smooth function f defined by (4.5) it is not smoothly solvable.

4.3. Choosing roots of polynomials differentiably

Here we use the results obtained in section 4.2 in order to construct global smooth roots, if a certain genericity condition ((1) or equivalently (2) in theorem 4.3.1) is fulfilled, and global differentiable roots always. The obstructions contained in the mentioned genericity condition arise in a natural way from the algorithm 4.2.9.

4.3.1. THEOREM. *Consider a smooth curve of hyperbolic polynomials*

$$P(t)(x) = x^n + a_2(t)x^{n-2} - \cdots + (-1)^n a_n(t) \qquad (t \in \mathbb{R}).$$

Let one of the following equivalent conditions be satisfied:

(1) *If two of the increasingly ordered continuous roots meet of infinite order somewhere, then they are equal everywhere.*
(2) *Let k be maximal with the property that $\tilde{\Delta}_k(P)$, from section 3.1, does not vanish identically for all t. Then $\tilde{\Delta}_k(P)$ vanishes nowhere of infinite order.*

Then the roots of P can be chosen smoothly, and any two choices differ by a permutation of the roots.

PROOF. *The local situation.* We claim that for any t_0, without loss $t_0 = 0$, the following conditions are equivalent:

(1) If two of the increasingly ordered continuous roots meet of infinite order at $t = 0$, then their germs at $t = 0$ are equal.
(2) Let k be maximal with the property that the germ at $t = 0$ of $\tilde{\Delta}_k(P)$ is not 0. Then $\tilde{\Delta}_k(P)$ is not infinitely flat at $t = 0$.
(3) The algorithm 4.2.9 never leads to step (3c).

To the proof of the claim:

(3) $\Rightarrow$ (1): Suppose for contradiction that two of the increasingly ordered continuous roots with different germs at $t = 0$ meet of infinite order at $t = 0$. Then in each application of step (2) in algorithm 4.2.9 these two roots stay with the same factor. After any application of step (3a) these two roots lead to roots with different germs at $t = 0$ of the modified polynomial which still meet of infinite order at $t = 0$. Hence, they never end up in a factor leading to step (3b) or to step (1). Since the algorithm has to stop after finitely many steps and step (3c) is the only remaining exit, the two roots end up in a factor leading to step (3c), a contradiction.

(1) $\Rightarrow$ (2): Let $x_1(t) \le \cdots \le x_n(t)$ be the continuous roots of $P(t)$, and let k be as required in (2). From (3.3) we have

$$\tilde{\Delta}_k(P) = \sum_{i_1 < i_2 < \cdots < i_k} (x_{i_1} - x_{i_2})^2 \cdots (x_{i_1} - x_{i_k})^2 \cdots (x_{i_{k-1}} - x_{i_k})^2.$$

Since the germ at $t = 0$ of $\tilde{\Delta}_k(P)$ is not 0, the germ at $t = 0$ of one summand in the above formula is not 0. If $\tilde{\Delta}_k(P)$ were infinitely flat at $t = 0$, then each summand had to be infinitely flat at $t = 0$, and, consequently, there had to be two roots among the x_i appearing in this summand which met of infinite order. By assumption their germs at $t = 0$ were equal, so each summand and thus $\tilde{\Delta}_k(P)$ vanished identically near $t = 0$, a contradiction.

(2) $\Rightarrow$ (3): Let k be as required in (2). Since $\tilde{\Delta}_k(P)$ vanishes only of finite order at $t = 0$, P has exactly k different roots off 0. Assume for contradiction that the algorithm 4.2.9 leads to step (3c), then $P = P^{(\infty)} \cdot P^{(s)}$ for a nontrivial polynomial $P^{(\infty)}$. Let $x_1(t) \le \cdots \le x_p(t)$ be the roots of $P^{(\infty)}(t)$ and $x_{p+1}(t) \le \cdots \le x_n(t)$ those of $P^{(s)}(t)$. We know that each x_i meets some x_j of infinite order and does not meet any x_l of infinite order, for $1 \le i, j \le p < l \le n$. Denote by $k^{(\infty)}$ and $k^{(s)}$ the number of generically different roots of $P^{(\infty)}$ and $P^{(s)}$, respectively, then $k^{(\infty)} \ge 2$

and $k = k^{(\infty)} + k^{(s)}$. Now, the only summand in the above formula for $\tilde{\Delta}_k(P)$ that does not vanish identically near 0 is the one involving exactly the k generically different roots of P near 0. Hence, this summand involves exactly $k^{(\infty)}$ generically different roots from $P^{(\infty)}$. But then we find two of them which meet each other of infinite order at 0, whence $\tilde{\Delta}_k(P)$ vanishes of infinite order at 0, contradicting the assumptions.

The global situation. From the first part of the proof we see that the algorithm 4.2.9 allows to choose the roots of P smoothly in a neighborhood of each point $t \in \mathbb{R}$ and that any two choices differ by a (constant) permutation of the roots. Thus we may glue the local solutions to a global solution. This completes the proof. $\square$

4.3.2. THEOREM. *Consider a curve of hyperbolic polynomials*

$$P(t)(x) = x^n - a_1(t)x^{n-1} + \cdots + (-1)^n a_n(t) \qquad (t \in \mathbb{R}),$$

where a_k is of class C^n for all $1 \le k \le n$. Then, there is a differentiable curve $x = (x_1, \dots, x_n) : \mathbb{R} \to \mathbb{R}^n$ whose components parameterize the roots of P.

PROOF. We follow one step of the algorithm 4.2.9. Without loss of generality we may assume that $a_1 = 0$: Replace x by $y = x - \frac{a_1(t)}{n}$ and note that a_1 is C^n. We want to prove first that there is a choice of differentiable roots locally near every $t \in \mathbb{R}$. So let $t_0 \in \mathbb{R}$ be arbitrary but fixed. Without loss of generality we may assume that $t_0 = 0$.

If $a_2(0) = 0$, then a_2 vanishes of second order at 0, since $\tilde{\Delta}_2(P(t)) = -2na_2(t)$ is non-negative, by theorem 3.1.2. Thus $a_2(t) = t^2 a_{2,2}(t)$ near 0 for a continuous function $a_{2,2}$. By the strong multiplicity lemma 4.2.8, we have $a_k(t) = t^k a_{k,k}(t)$ near 0 for continuous functions $a_{k,k}$, for all $2 \le k \le n$. Let us consider the following continuous curve of polynomials

$$P_1(t)(x) := x^n + a_{2,2}(t)x^{n-2} - a_{3,3}(t)x^{n-3} + \cdots + (-1)^n a_{n,n}(t).$$

Note that $P_1(t)(x) = t^{-n}P(t)(tx)$. It follows that $P_1(t)$ is hyperbolic for all t. Let $z_1(t) \le \cdots \le z_n(t)$ be its continuous roots, by theorem 2.3.2. Then, $x_j(t) := t \cdot z_j(t)$, where $1 \le j \le n$, parameterize the roots of $P(t)$, and they are differentiable at 0:

$$\lim_{t \to 0} \frac{t \cdot z_j(t)}{t} = \lim_{t \to 0} z_j(t) = z_j(0).$$

However, note that $x_j(t) = y_j(t)$ for $t \ge 0$, but $x_j(t) = y_{n-j}(t)$ for $t \le 0$, where $y_1(t) \le \cdots \le y_n(t)$ are the ordered continuous roots of $P(t)$. This gives us one choice of differentiable roots near $t = 0$. Any such choice is then given by this choice and applying afterwards any permutation of the set $\{1, \dots, n\}$ keeping invariant the function $j \mapsto z_j(0)$, i.e., keeping invariant the derivatives at 0 of the roots.

If $a_2(0) \ne 0$, then not all roots of $P(0)$ are equal. By the splitting lemma 4.2.3, we may factorize $P(t) = P_1(t) \cdots P_l(t)$, where the coefficients of the $P_i(t)$ have the differentiability conditions required in the theorem and where each $P_i(0)$ has all roots equal to c_i with pairwise distinct c_i. But then we can treat each P_i separately, and for each P_i the previous case occurs. Therefore, the roots of each P_i and hence of P can be arranged differentiably near $t = 0$.

Note that we have to apply a permutation on one side of 0 to the original roots, in the following case: Two roots x_k and x_l meet at 0 slowly, i.e., $x_k(t) - x_l(t) = t \cdot c_{kl}(t)$ with $c_{kl}(0) \ne 0$ which means that their derivatives at 0 disagree. We may apply to this choice an arbitrary permutation of any two roots x_k and x_l which meet with $c_{kl}(0) = 0$ (i.e. at least of second order), and we get thus any differentiable choice of roots near $t = 0$.

Now let us construct global differentiable roots of P out from the local ones existing near any $t \in \mathbb{R}$. We start with the increasingly ordered continuous roots

$y_1(t) \le \cdots \le y_n(t)$. Then we put

$$x_j(t) = y_{\sigma(t)(j)}(t) \qquad (1 \le j \le n),$$

where the permutation $\sigma(t)$ is given by

$$\sigma(t) = (1,2)^{\epsilon_{1,2}(t)} \dots (1,n)^{\epsilon_{1,n}(t)}(2,3)^{\epsilon_{2,3}(t)} \dots (n-1,n)^{\epsilon_{n-1,n}(t)},$$

and where $\epsilon_{i,j}(t) \in \{0,1\}$ will be specified as follows: On the closed set $S_{i,j}$ of all t, where $y_i(t)$ and $y_j(t)$ meet at least of second order any choice is good. The complement of $S_{i,j}$ in $\mathbb{R}$ is an at most countable union of open intervals. In each interval we choose a point, where we put $\epsilon_{i,j}(t) = 0$. Going right (and left) from this point we change $\epsilon_{i,j}(t)$ in each point, where y_i and y_j meet slowly. Since these points accumulate only in $S_{i,j}$, this construction is well-defined and leads to a global differentiable parameterization of the roots of P. □

Remark. Note that the following statement is due to Bronshtein [**Bro79**], compare with theorem 5.2.1: *For any $t_0 \in \mathbb{R}$ there exist n roots $x_j(t_0;t)$ $(1 \le j \le n)$ of $P(t)$ such that $x_j(t_0;t)$ are differentiable at $t = t_0$.* The differentiability of $x_j(t_0;t)$ is assured only at $t = t_0$. The proof of existence of global differentiable roots of P is due to Mandai [**Man85**].

4.4. The real analytic case

The algorithm 4.2.9 motivates in a natural way to consider real analytic curves of hyperbolic polynomials and investigate their solvability, since in the real analytic case step (3c) in the algorithm 4.2.9 cannot occur.

So let $P(t)(x) = x^n - a_1(t)x^{n-1} + \cdots + (-1)^n a_n(t)$ be a curve of hyperbolic polynomials, where all $a_i(t)$ are real analytic in t. Recall from definition 4.2.1, that we say that P is real analytically solvable, if we may find functions $x_i(t)$ for $i = 1, \dots, n$ which are real analytic in t and parameterize the roots of $P(t)$ for all t.

4.4.1. Theorem. *Let P be a real analytic curve of hyperbolic polynomials*

$$P(t)(x) = x^n - a_1(t)x^{n-1} + \cdots + (-1)^n a_n(t) \qquad (t \in \mathbb{R}).$$

Then P is real analytically solvable, globally on $\mathbb{R}$. All solutions differ by permutations.

Proof. We first show that P is real analytically solvable, locally near each point $t_0 \in \mathbb{R}$. Without loss of generality we may assume that $t_0 = 0$. Furthermore, we can suppose without loss that $a_1 = 0$.

The proof will be carried out by induction on the polynomial degree n. If $n = 1$, then the theorem evidently holds. Let us assume that the statement is true for degrees strictly smaller than $n > 1$. We consider several cases:

The case $a_2(0) \neq 0$. Here not all roots of $P(0)$ are equal, so by the splitting lemma 4.2.3 we may factorize $P(t) = P_1(t) \cdot P_2(t)$ for real analytic curves of hyperbolic polynomials, P_1 and P_2, of positive degree. Hence we have reduced the problem to lower degree, whence by induction hypothesis we find a real analytic choice of roots near 0.

The case $a_2(0) = 0$. If $a_2 = 0$, then by lemma 4.2.6 all roots of P are identically 0 and we are done. Otherwise, for the multiplicity of the real analytic function a_2 at 0 we have $1 \le m(a_2) < \infty$, and, again by lemma 4.2.6, all roots of $P(0)$ are 0. The multiplicity of a_2 at 0 cannot be odd, for otherwise $\tilde{\Delta}_2(P)(t) = -2na_2(t)$ changed sign at $t = 0$ contradicting the hyperbolicity of P, according to theorem 3.1.2. So we write $m(a_2) = 2r$ for a positive integer r. Then by the multiplicity lemma 4.2.7 we have $a_k(t) = t^{kr} a_{k,kr}(t)$ for real analytic $a_{k,kr}$ for all $2 \le k \le n$. Let us consider the following real analytic curve of hyperbolic polynomials

$$P_r(t)(x) = x^n + a_{2,2r}(t)x^{n-2} - a_{3,3r}(t)x^{n-3} + \cdots + (-1)^n a_{n,nr}(t).$$

Note that, if $P_r(t)$ is real analytically solvable and $x_j(t)$, for $1 \le j \le n$, are its real analytic roots, then $t^r x_j(t)$, for $1 \le j \le n$, are the roots of $P(t)$ and so the original curve P is real analytically solvable, too. Now $a_{2,2r}(0) \ne 0$ and we are done by the case above. This completes the proof of local real analytic solvability.

Now let $x = (x_1, \ldots, x_n) : I \to \mathbb{R}^n$ be a real analytic curve of roots of P on an open interval $I \subseteq \mathbb{R}$. Then we assert that any real analytic curve of roots of P on I is of the form $\alpha \circ x$ for some permutation α. Let $y : I \to \mathbb{R}^n$ be another real analytic curve of roots of P. Let $t_k \to t_0$ be a convergent sequence of distinct points in I. Then $y(t_k) = \alpha_k(x(t_k)) = (x_{\alpha_k(1)}(t_k), \ldots, x_{\alpha_k(n)}(t_k))$ for permutations α_k. By choosing a subsequence of (t_k) we may assume that all α_k are the same permutation α. But then the real analytic curves y and $\alpha \circ x$ coincide on a converging sequence, so they coincide on whole I and the assertion follows.

The local real analytic solvability and the uniqueness of real analytic solutions up to permutations, we have shown, suffice to glue a global real analytic parameterization of the roots of P on entire $\mathbb{R}$. This completes the proof. □

4.4.2. Remarks. (1) Note that the local existence part of this theorem is due to Rellich [**Rel37**, Hilfssatz 2]. His proof uses Puiseux-expansions.

(2) The uniqueness statement of theorem 4.4.1 is wrong in the smooth case (without restrictions on the roots), as is shown by the following example: $x^2 = f(t)^2$, where f is smooth. In each point t where f is infinitely flat one can change sign in the solution $x(t) = \pm f(t)$ without destroying its smoothness. No sign change can be absorbed in a permutation (constant in t). If there are infinitely many points of flatness for f we get uncountably many smooth solutions.

(3) Theorem 4.4.1 reminds of the curve lifting property of covering mappings. But unfortunately one cannot lift real analytic homotopies, as the following example shows. This example also shows that polynomials which are real analytically parameterized by higher dimensional variables are not real analytically solvable. Consider the 2-parameter family $x^2 = t_1^2 + t_2^2$. The two continuous solutions are $x(t) = \pm|t|$ with $t = (t_1, t_2)$, but for none of them $t_1 \mapsto x(t_1, 0)$ is differentiable at 0. There remains the question whether for a real analytic submanifold of the space of hyperbolic polynomials one can choose the roots real analytically along this manifold. This is not the case: Consider

$$P(t_1, t_2)(x) = (x^2 - (t_1^2 + t_2^2))(x - (t_1 - a_1))(x - (t_2 - a_2)),$$

which is not real analytically solvable by the above arguments. For $a_1 \ne a_2$ the coefficients describe a real analytic embedding for (t_1, t_2) near 0.

4.5. The ultradifferentiable case

We consider the case when the coefficients of the hyperbolic polynomial

$$P(t)(x) = x^n - a_1(t)x^{n-1} + \cdots + (-1)^n a_n(t)$$

belong to a ultradifferentiable class $C^M(I)$, where I is an open interval. We shall show that the roots of P may be parameterized by functions which are in the class $C^M(I)$ as well, if $C^M(I)$ is quasi-analytic. This is due to Chaumat and Chollet [**CC04**].

4.5.1. Ultradifferentiable functions. Let $M = (M_k)_{k\in\mathbb{N}_0}$ be an increasing sequence of positive reals and let $U \subseteq \mathbb{R}^m$ be open. A C^∞ function f defined on U belongs to the *ultradifferentiable* class $C^M(U)$, if for each compact $K \subseteq U$ there exists a constant $C_K \geq 1$ such that

$$\sup_{\substack{x\in K\\ \alpha\in\mathbb{N}_0^m,|\alpha|=l}} \frac{|\partial^\alpha f(x)|}{l! M_l C_K^l} < \infty,$$

where $\partial^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_m}\right)^{\alpha_m}$.

Moreover, we will impose additional conditions on M:

(D_1) $M_0 \geq 1$ and $(M_k)_{k\in\mathbb{N}_0}$ is *logarithmically convex*, i.e., the sequence $\left(\frac{M_{k+1}}{M_k}\right)_{k\in\mathbb{N}_0}$ is increasing.

(D_2) $\frac{M_{k+1}}{M_k} \to \infty$.

(D_3) There exists some $A \geq 1$ such that $M_{k+1} \leq A^{k+1} M_k$ for all $k \geq 0$.

Condition (D_1) guarantees that the class $C^M(U)$ is invariant under composition. (D_2) assures that $C^M(U)$ contains more than just the analytic functions. If condition (D_3) is fulfilled, the class $C^M(U)$ is stable under differentiation.

The class $C^M(U)$ is called *quasi-analytic* if and only if

$$\sum_{k\geq 1} \frac{1}{k M_k^{1/k}} = \infty.$$

By the Denjoy-Carleman theorem, this is equivalent to the following condition: If $f \in C^M(U)$ and $\partial^\alpha f(0) = 0$ for all $\alpha \in N_0^m$, then $f = 0$. For more on quasi-analytic functions see e.g. [**Hör83a**] and [**Rud87**].

4.5.2. Lemma. *Let M be an increasing sequence of positive reals satisfying (D_1), (D_2), and (D_3). Let $U \subseteq \mathbb{R}^m$ be open and connected and assume that $0 \in U$. Consider a monic polynomial*

$$P(t)(z) = z^n - a_1(t)z^{n-1} + \cdots + (-1)^n a_n(t)$$

whose coefficients a_i belong to the class $C^M(U)$. If $P(0)(z) = z^s Q(z)$ with $Q(0) \neq 0$, then there exists an open and connected subset $U' \in \mathbb{R}^m$ with $0 \in U' \subseteq U$ and there exist two monic polynomials P_1 and P_2 of degree s and $n-s$, respectively, whose coefficients belong to $C^M(U')$ such that $P(t) = P_1(t)P_2(t)$, $P_1(0)(z) = z^s$, and $P_2(0)(z) = Q(z)$.

Proof. Since P is holomorphic in z and continuous in t, there exists some $r > 0$ and an open neighborhood U' of 0 with $U' \subseteq U$ such that for all $t \in U'$ the polynomial $P(t)$ has exactly s roots $\lambda_1(t), \ldots, \lambda_s(t)$ (counted with multiplicities) in the disk centered at 0 with radius r and no root on its boundary, see theorem 2.2.3. By the theorem of residues,

$$\frac{1}{2i\pi}\int_{|\zeta|=r} \frac{\partial P(t)(\zeta)}{\partial z}\frac{1}{P(t)(\zeta)}\zeta^k d\zeta = \sum_{j=1}^s \lambda_j(t)^k =: b_k(t) \qquad (1 \leq k \leq s).$$

The functions $b_1, \ldots, b_s$ belong to $C^M(U')$ and, by (3.1), also the coefficients of the polynomial $P_1(t)(z) := \prod_{j=1}^s (z - \lambda_j(t))$ are in $C^M(U')$.

The quotient $\frac{P}{P_1}$ is a holomorphic function in the disk $|z| < r$ which does not vanish on the boundary $|z| = r$, for all $t \in U'$. Evidently $\frac{P}{P_1}$ is a polynomial. It is easy to see from Cauchy's formula applied to the boundary of the disk $|z| < r$ that the mapping $t \mapsto \frac{P}{P_1}(t)(z) =: P_2(t)(z)$ is in $C^M(U')$. This completes the proof. $\square$

4.5.3. THEOREM. *Let M be an increasing sequence of positive reals satisfying (D_1), (D_2), and (D_3). Let $I \subseteq \mathbb{R}$ be an open interval. Consider a monic hyperbolic polynomial*

$$P(t)(z) = z^n - a_1(t)z^{n-1} + \cdots + (-1)^n a_n(t)$$

whose coefficients a_i belong to the quasi-analytic class $C^M(I)$. Then there exist functions $\lambda_1(t), \ldots, \lambda_n(t)$ in the quasi-analytic class $C^M(I)$ which parameterize the roots of $P(t)$ for each $t \in I$.

PROOF. We first show that the roots of P may be parameterized by functions in C^M, locally near each point $t_0 \in \mathbb{R}$. Without loss of generality we may assume that $t_0 = 0$. Furthermore, we can suppose without loss that $a_1 = 0$.

The proof will be carried out by induction on n. If $n = 1$, then the statement is evident. Let us assume that the statement is true for degrees strictly smaller than $n > 1$. We consider two cases:

The case $a_2(0) \neq 0$. Here not all roots of $P(0)$ are equal. Without loss we can assume that $P(0)(z) = z^s Q(z)$ with $Q(0) \neq 0$ and $1 \leq s < n$. By lemma 4.5.2 we may write $P(t) = P_1(t) \cdot P_2(t)$ on an interval $0 \in J \subseteq I$ for hyperbolic polynomials P_1 and P_2 whose coefficients belong to $C^M(J)$. Hence we have reduced the problem to lower degree, whence by induction hypothesis we find a parameterization of the roots of P by functions in $C^M(J)$.

The case $a_2(0) = 0$. If $a_2 = 0$, then by lemma 4.2.6 all roots of P are identically 0 and we are done.

Otherwise, by quasi-analyticity, for the multiplicity of the function a_2 at 0 we have $1 \leq m(a_2) = 2r < \infty$, where r is a positive integer. The multiplicity lemma 4.2.7 remains valid in the quasi-analytic setting with the same proof. We conclude that $a_k(t) = t^{kr} a_{k,kr}(t)$ for $a_{k,kr}$ in $C^M(J)$, for all $2 \leq k \leq n$ and for an interval $0 \in J \subseteq I$. Let us consider the following quasi-analytic curve of hyperbolic polynomials

$$P_r(t)(x) = x^n + a_{2,2r}(t)x^{n-2} - a_{3,3r}(t)x^{n-3} + \cdots + (-1)^n a_{n,nr}(t).$$

Note that, if the roots of $P_r(t)$ can be parameterized by functions $\mu_j(t)$, for $1 \leq j \leq n$, of class C^M then $t^r \mu_j(t)$, for $1 \leq j \leq n$, provide a parameterization of class C^M of the roots of $P(t)$. Now $a_{2,2r}(0) \neq 0$ and we are done by the case above. This completes the proof of the local statement.

Quasianalyticity implies analogously to the real analytic case in 4.4.1 that the local parameterizations may be glued to a global parameterization of the roots of P in the class C^M. □

We shall see in section 7.1 that a C^{3n} curve of monic hyperbolic polynomials of degree n allows a twice differentiable parameterization of its roots. By 4.1.4 and 4.1.5, we cannot expect more than twice differentiability of the roots even starting with a C^∞ curve of hyperbolic polynomials. However, one might ask whether there exists an ultradifferentiable non-quasi-analytic class $C^M(\mathbb{R})$ such that the roots of a monic hyperbolic polynomial with coefficients in $C^M(\mathbb{R})$ can be chosen more regular than twice differentiable. The following example which is a modification of an example in [**BBCP06**] shows that the answer is no.

4.5.4. EXAMPLE. Let M be an increasing sequence of positive reals satisfying (D_1), (D_2), and (D_3). *Given any modulus of continuity ω, there exists a non-negative function f in the non-quasi-analytic class $C^M(\mathbb{R})$ which vanishes only at 0 and is infinitely flat at 0 such that $h = (\sqrt{f})'$ is not ω-continuous.* Since $C^M(\mathbb{R})$ is not quasi-analytic, there exists a function $\chi \in C^M(\mathbb{R})$ vanishing outside $(-2, 2)$, positive on $(-2,2)$, and such that $\chi|_{[-1,1]} = 1$ (e.g. [**Rud87**]). For $n \geq 1$ let

$$\rho_n = \frac{1}{n^2}, \quad t_n = 2\rho_n + \sum_{j=n+1}^{\infty} 3\rho_j,$$
$$I_n = [t_n - \rho_n, t_n + \rho_n], \quad J_n = [t_{n+1} + \rho_{n+1}, t_n - \rho_n],$$
$$\alpha_n = \frac{1}{2^n}, \quad \epsilon_n = \omega^{-1}(\frac{\alpha_n}{2}), \quad \beta_n = \alpha_n \epsilon_n^2,$$

where ω^{-1} is the inverse function of ω. It is no restriction to assume that $\omega(s) \geq s$ for all s. Hence $\epsilon_n \leq \frac{\alpha_n}{2} \leq \rho_n$ and $t_n + \epsilon_n \in I_n$. The function

$$f(t) = \begin{cases} \chi^2\left(-2^{\frac{t_1+2\rho_1-t}{\rho_1}}\right) + \sum_{n=1}^{\infty} \chi^2\left(\frac{t-t_n}{\rho_n}\right)(\alpha_n(t-t_n)^2 + \beta_n) & \text{if } t \geq 0 \\ f(-t) & \text{if } t < 0 \end{cases}$$

belongs to $C^M(\mathbb{R})$ and is strictly positive for $t \neq 0$, but $h = (\sqrt{f})'$ is not ω-continuous e.g. on $[-1,1]$. Indeed, it is easy to estimate

$$|f^{(k)}(t)| \leq C_k \alpha_n \rho_n^{-k} \to \infty, \quad \text{as } n \to \infty,$$

for $t \in J_n \cup I_n$, which shows that $f \in C^\infty(\mathbb{R})$. The sum in the definition of f consists of finitely many summands for each t. Since $C^M(\mathbb{R})$ is an algebra and also invariant under composition, we conclude that $f \in C^M(\mathbb{R})$. However, we have

$$\frac{|h(t_n+\epsilon_n) - h(t_n)|}{\omega(\epsilon_n)} = \frac{\alpha_n \epsilon_n}{\sqrt{\alpha_n \epsilon_n^2 + \beta_n}\,\omega(\epsilon_n)} = \frac{\sqrt{a_n}}{\sqrt{2}\omega(\epsilon_n)} = \frac{\sqrt{2}}{\sqrt{\alpha_n}}$$

which goes to infinity as $n \to \infty$.

4.6. The complex case

In this section we shall investigate the solvability of curves of polynomials

$$P(t)(z) = z^n - a_1(t)z^{n-1} + \cdots + (-1)a_n(t)$$

with complex valued coefficients $a_1(t), \ldots, a_n(t)$. In particular, we shall study the problem of finding smooth or real analytic curves of complex roots for smooth or real analytic curves $t \mapsto P(t)$, respectively, for real parameter t, and we shall investigate the holomorphic case when t is complex and $P(t)$ is holomorphic in t.

4.6.1. The quadratic case. As in the hyperbolic situation we will start with discussing the case $n = 2$. Let f be a smooth complex valued function, defined near $0 \in \mathbb{R}$, such that $f(0) = 0$. We look for a smooth complex valued function g, defined near $0 \in \mathbb{R}$, with $f = g^2$. If $m(f)$ is finite and even, then we have $f(t) = t^{m(f)} h(t)$ with smooth h satisfying $h(0) \neq 0$, and $g(t) := t^{\frac{m(f)}{2}} \sqrt{h(t)}$ is a local solution. If $m(f)$ is finite and odd, there is no smooth solution g, also not in the real analytic and holomorphic cases. If instead $f(t)$ is flat at $t = 0$, then one has no definite answer, and for the concrete f given in equation (4.5) there still not exists a smooth square root.

4.6.2. Note that the preliminaries presented in section 3.1, including the definition of the Bezoutiant B, its principal minors Δ_k, and formula (3.3), are still valid in the present case, where now coefficients and roots are complex valued. But there are no restrictions on the coefficients, whence the space of monic polynomials of degree n may be identified with $\mathbb{C}^n$. Moreover, proposition 4.2.2 and the splitting lemma 4.2.3 are true in the complex case. Also proposition 4.2.5 remains valid, since it follows from (3.3). Evidently, lemma 4.2.6 does not hold anymore, and the multiplicity lemma 4.2.7 remains valid only partially:

Lemma (Multiplicity Lemma). *Consider a smooth (real analytic, holomorphic) curve of complex polynomials*

$$P(t)(z) = z^n + a_2(t)z^{n-2} - \cdots + (-1)^n a_n(t).$$

Then, for integers r, the following conditions are equivalent:

(1) $m(a_k) \geq kr$, *for all* $2 \leq k \leq n$;
(2) $m(\tilde{\Delta}_k) \geq k(k-1)r$, *for all* $2 \leq k \leq n$.

Proof. Without loss of generality we can assume $r > 0$.

(1) $\Rightarrow$ (2): Exactly the same arguments as in the proof of the multiplicity lemma 4.2.7 work.

(2) $\Rightarrow$ (1) : Since $\tilde{\Delta}_2 = -2na_2$ and $\tilde{s}_2 = -2a_2$, we find that $\tilde{s}_2(0) = 0$. Consequently, $\tilde{\Delta}_3(0) = -n\tilde{s}_3(0)^2$ and thus $\tilde{s}_3(0) = 0$. Going on like this we obtain $\tilde{s}_4(0) = \cdots = \tilde{s}_n(0) = 0$. Then by (3.1) we have $a_k(0) = 0$ for all $2 \leq k \leq n$. Now follow the proof of the multiplicity lemma 4.2.7((3) $\Rightarrow$ (1)). Then for the polynomial $\bar{P}_m$ from (4.8) we obtain $\tilde{\Delta}_k(\bar{P}_m(t)) = t^{-k(k-1)m}\tilde{\Delta}_k(P(t))$ and thus $m(\tilde{\Delta}_k(\bar{P}_m)) \geq k(k-1)(r-m)$, for all $2 \leq k \leq n$. Now $r - m > 0$, by (4.7), and so we may conclude as before that all coefficients of $\bar{P}_m(t)$ vanish for $t = 0$. But this is a contradiction for those k with $m_k = km$. $\square$

The proof shows that the multiplicity lemma 4.2.7 holds only partially by the lack of lemma 4.2.6.

4.6.3. Factorizing algorithm. As in section 4.2 we may construct an algorithm which extracts the solvable part from the original curve P:

Algorithm. Consider a smooth (real analytic, holomorphic) curve of polynomials

$$P(t)(z) = z^n - a_1(t)z^{n-1} + a_2(t)z^{n-2} - \cdots + (-1)^n a_n(t)$$

with complex coefficients. The algorithm has the following steps:

(1) If all roots of $P(0)$ are pairwise different, then P is smoothly (real analytically, holomorphically) solvable for t near 0 by proposition 4.2.2.
(2) If there are distinct roots at $t = 0$, we put them into two disjoint subsets which splits $P(t) = P_1(t) \cdot P_2(t)$ near 0 by the splitting lemma 4.2.3. We then feed $P_1(t)$ and $P_2(t)$ (which have lower degrees) into the algorithm.
(3) If all roots of $P(0)$ are equal, then we first reduce $P(t)$ to the case $a_1 = 0$ by replacing the variable z with $y = z - \frac{a_1(t)}{n}$. Then, by Vieta's formula for a_1, all roots of $P(0)$ are equal to 0. Consequently, $a_k(0) = 0$ for all $1 \leq k \leq n$.
(3a) If there does not exist an integer $r > 0$ with $m(a_k) \geq kr$ for $2 \leq k \leq n$, then the curve of polynomials P is not smoothly (real analytically, holomorphically) solvable, by proposition 4.2.5. We store the polynomial as an output of the procedure, as a factor of $P^{(n)}$ below.
(3b) If there exists an integer $r > 0$ with $m(a_k) \geq kr$ for $2 \leq k \leq n$ but not all $m(a_k)$ are infinite, write $a_k(t) = t^{kr}a_{k,kr}(t)$ for smooth (real analytic, holomorphic) $a_{k,kr}$ and $2 \leq k \leq n$. Let us consider the following smooth (real analytic, holomorphic) curve of polynomials

$$P_r(t)(z) := z^n + a_{2,2r}(t)z^{n-2} - a_{3,3r}(t)z^{n-3} + \cdots + (-1)^n a_{n,nr}(t).$$

If $P_r(t)$ is smoothly (real analytically, holomorphically) solvable and $z_j(t)$ are its smooth (real analytic, holomorphic) roots, then $t^r z_j(t)$ are the roots of $P(t)$ and hence the original curve P is smoothly (real analytically, holomorphically) solvable, too.

(3b.1) If for one coefficient a_k we have $m(a_k) = kr$, then $P_r(t)$ has a coefficient $a_{k,kr}$ which does not vanish at 0. So not all roots of $P_r(0)$ are equal, and we may feed P_r into step (2).

(3b.2) If all coefficients of $P_r(0)$ are zero, we feed P_r again into step (3).

(3c) In the smooth case all $m(a_k)$ can be infinite. Then we store the polynomial as a factor of $P^{(\infty)}$ below.

In the real analytic and holomorphic cases the algorithm provides a splitting of the original curve $P(t) = P^{(n)}(t) \cdot P^{(s)}(t)$ into real analytic and holomorphic curves, respectively, where $P^{(s)}$ is solvable and where $P^{(n)}$ is not solvable. But it may contain solvable roots.

In the smooth case the algorithm yields a factorization near $t = 0$ into smooth curves of polynomials $P(t) = P^{(\infty)}(t) \cdot P^{(n)}(t) \cdot P^{(s)}(t)$, where $P^{(\infty)}$ has the property that each root meets another one of infinite order at $t = 0$, where $P^{(s)}$ is smoothly solvable, and no two roots meet of infinite order at 0, and where $P^{(n)}$ is not smoothly solvable but may contain solvable roots.

4.6.4. Remark. If $P(t)$ is a polynomial whose coefficients are meromorphic functions of a complex variable t, there is a well developed theory of the roots of $P(t)(z) = 0$ as multi-valued meromorphic functions, given by Puiseux or Laurent-Puiseux series. But it is difficult to extract holomorphic information out of it. See for example Theorem 3 on page 370 (Anhang, §5) of [**Bau72**].

4.6.5. Remark. In the absence of hyperbolicity we cannot expect that the roots of a polynomial with smooth coefficients are Lipschitz continuous, however they may possess a weaker regularity: They may be absolute continuous, which still allows us to differentiate them almost everywhere. For instance, the roots of the equation $z^2 = t$ cannot be Lipschitz continuous near $t = 0$, but

$$z_1(t) = \begin{cases} \sqrt{t} & \text{for} \quad t \geq 0 \\ \sqrt{-t}i & \text{for} \quad t < 0 \end{cases} \quad \text{and} \quad z_2(t) = \begin{cases} -\sqrt{t} & \text{for} \quad t \geq 0 \\ -\sqrt{-t}i & \text{for} \quad t < 0 \end{cases}$$

gives an absolute continuous parameterization of its roots.

In fact it has been proved by Spagnolo [**Spa99**] that for every complex polynomial of degree $n \leq 3$, with coefficients depending smoothly on a real parameter t, it is possible to select a n-tuple of roots absolutely continuous in t. For arbitrary degree n this question seems to be unsolved.

CHAPTER 5

Bronshtein's approach

In this chapter we consider Bronshtein's approach who already in 1979 could prove that the roots of a C^n curve of hyperbolic polynomials of degree n may be chosen differentiable with locally bounded derivatives. The whole chapter is based on [**Bro79**].

Note that in this chapter we will use several times Taylor's formula without specifying the intermediate values $\xi, \eta, \dots$ in the remainder term. They are always to be interpreted in the obvious way.

5.1. Introduction: degree 3

To get an idea, how Bronshtein proves the local boundedness of the derivatives of roots of hyperbolic polynomials depending smoothly on a parameter, we want to discuss here the case when the polynomials have degree 3. It shortens and simplifies the general proof essentially but uses its whole machinery of argumentation.

This discussion includes the treatment of the quadratic case. So the reader may compare it to Alekseevsky, Kriegl, Losik, and Michor's consideration of this case in proposition 4.1.1.

5.1.1. Let us consider a curve of monic hyperbolic polynomials of degree 3

$$P(t)(x) = x^3 + A_1(t)x^2 + A_2(t)x + A_3(t),$$

where $t \in [-1,1]$. Assume that the coefficients satisfy $A_i \in C^i([-1,1])$, for $1 \le i \le 3$, $A_2(0) \neq 0$, and $A_3(0) = 0$. *We want to show that there exists a positive constant C such that*

$$\left|\frac{A_3'(0)}{A_2(0)}\right| \le \left(\max_{i,t,j\le i} |A_i^{(j)}(t)| + 2\right)^C.$$

To shorten notation let us introduce $a_i = A_i(0)$, for $1 \le i \le 3$. We define

$$M_0 = 27\left(\max_{i,t,j\le i} |A_i^{(j)}(t)| + 4\right)^6 \qquad \text{and} \qquad M_i = M_{i-1}^{100} \quad (1 \le i \le 3).$$

Their exact value is not really important, what counts is that they are chosen large enough for the estimates to come.

Since the polynomial $P(t)$ has only real roots, the same is true for $\frac{\partial}{\partial x}P(t)(x) = 3x^2 + 2A_1(t)x + A_2(t)$, see lemma 5.4.4. And this is equivalent to

$$A_2(t) \le \frac{1}{3}A_1^2(t). \tag{5.1}$$

Consider the following two cases separately:

$$(A): \ |a_2| \le a_1^2 \qquad \text{and} \qquad (B): \ |a_2| \ge a_1^2.$$

We start with case (A): The assumption $a_2 \neq 0$ implies $a_1 \neq 0$. Consider $\frac{A_1(t)}{a_1} = 1 + \frac{A_1^{(1)}(\xi)}{a_1}t$, for $|t| \le M_1^{-1}|a_1|$:

$$\frac{1}{2} \le 1 - \frac{|A_1^{(1)}(\xi)|}{|a_1|}|t| \le \frac{A_1(t)}{a_1} \le 1 + \frac{|A_1^{(1)}(\xi)|}{|a_1|}|t| \le 2, \tag{5.2}$$

since $\frac{|A_1^{(1)}(\xi)|}{|a_1|}|t| \leq M_0 M_1^{-1} \leq \frac{1}{2}$. Put $t = \pm M_0^{-1} a_1$ into (5.1) and use Taylor's formula:

$$a_2 \pm a_2^{(1)} M_0^{-1} a_1 + \frac{A_2^{(2)}(\xi)}{2!} M_0^{-2} a_1^2 \leq \frac{1}{3}(a_1 \pm A_1^{(1)}(\eta) M_0^{-1} a_1)^2.$$

Consequently,

$$\begin{aligned}
\pm a_2^{(1)} M_0^{-1} a_1 &\leq \frac{1}{3} a_1^2 \pm \frac{2}{3} A_1^{(1)}(\eta) M_0^{-1} a_1^2 + \frac{1}{3}(A_1^{(1)}(\eta))^2 M_0^{-2} a_1^2 \\
&\quad - a_2 - \frac{A_2^{(2)}(\xi)}{2!} M_0^{-2} a_1^2 \\
&\leq \frac{1}{3} a_1^2 + \frac{2}{3}|A_1^{(1)}(\eta)| M_0^{-1} a_1^2 + \frac{1}{3}(A_1^{(1)}(\eta))^2 M_0^{-2} a_1^2 \\
&\quad + |a_2| + \frac{|A_2^{(2)}(\xi)|}{2!} M_0^{-2} a_1^2 \\
&\leq M_0 a_1^2,
\end{aligned}$$

by the definition of M_0. Therefore, $|a_2^{(1)}| \leq M_0^2 |a_1|$ which gives, for $|t| \leq M_1^{-1}|a_1|$,

$$\begin{aligned}
|A_2(t)| = \left| a_2 + a_2^{(1)} t + \frac{A_2^{(2)}(\xi)}{2!} t^2 \right| &\leq |a_2| + |a_2^{(1)}||t| + \frac{|A_2^{(2)}(\xi)|}{2!}|t|^2 \\
&\leq a_1^2 + M_0^2 |a_1| M_1^{-1} |a_1| + M_0 M_1^{-2} a_1^2 \leq 2a_1^2.
\end{aligned} \tag{5.3}$$

For $|t| \leq M_1^{-1}|a_1|$, we get

$$\begin{aligned}
|A_2'(t)| = |a_2^{(1)} + A_2^{(2)}(\xi) t| &\leq |a_2^{(1)}| + |A_2^{(2)}(\xi)||t| \\
&\leq M_0^2 |a_1| + M_0 M_1^{-1} |a_1| \leq M_0^3 |a_1|,
\end{aligned}$$

whence, for $|t| \leq \frac{1}{2} M_1^{-1} \left|\frac{a_2}{a_1}\right|$ $\big(\leq M_1^{-1}|a_1|, \text{ by } (5.3)\big)$,

$$|A_2(t) - a_2| = |A_2'(\xi)||t| \leq M_0^3 |a_1| \frac{1}{2} M_1^{-1} \left|\frac{a_2}{a_1}\right| \leq \frac{1}{2}|a_2|,$$

implying

$$\frac{1}{2} \leq \frac{A_2(t)}{a_2} \leq 2, \tag{5.4}$$

for $|t| \leq M_2^{-1} \left|\frac{a_2}{a_1}\right|$.

For a root $x(t)$ of $P(t)$, the following estimate holds

$$|A_3(t)| \leq |x(t)|^3 + |A_1(t)||x(t)|^2 + |A_2(t)||x(t)|.$$

Let $x_1(t)$ and $x_2(t)$ be the roots of $\frac{\partial}{\partial x} P(t)(x)$ ordered such that $|x_1(t)| \leq |x_2(t)|$ for all t. By Vieta's formulas, $\frac{1}{3}|A_2(t)| = |x_1(t) x_2(t)|$ and $\frac{2}{3}|A_1(t)| = |x_1(t) + x_2(t)| \leq |x_1(t)| + |x_2(t)| \leq 2|x_2(t)|$. It implies that $|x_1(t)| = \frac{|x_1(t)x_2(t)|}{|x_2(t)|} \leq \left|\frac{A_2(t)}{A_1(t)}\right|$; if $x_2(t) = 0$, then $x_1(t) = 0$, and the inequality is trivial. So, if $|x_2(t)| \leq 4|x_1(t)|$, we have $|x_2(t)| \leq 4\left|\frac{A_2(t)}{A_1(t)}\right|$. And, if $|x_2(t)| > 4|x_1(t)|$, consider $\frac{2}{3}|A_1(t)| = |x_1(t) + x_2(t)| = \left|x_2(t)\left(1 + \frac{x_1(t)}{x_2(t)}\right)\right| = |x_2(t)|\left|1 + \frac{x_1(t)}{x_2(t)}\right| \geq |x_2(t)|\left|1 - \left|\frac{x_1(t)}{x_2(t)}\right|\right| \geq \frac{3}{4}|x_2(t)|$. Thus, in any case, we have

$$|x_2(t)| \leq 4\left(|A_1(t)| + \left|\frac{A_2(t)}{A_1(t)}\right|\right).$$

Since all roots of $P(t)$ are real, there has to be a root $x(t)$ of $P(t)$ lying between $x_1(t)$ and $x_2(t)$ (see lemma 5.4.4). Therefore,

$$\begin{aligned}|A_3(t)| \le 64\left(|A_1(t)| + \left|\frac{A_2(t)}{A_1(t)}\right|\right)^3 + 16|A_1(t)|\left(|A_1(t)| + \left|\frac{A_2(t)}{A_1(t)}\right|\right)^2 \\ + 4|A_2(t)|\left(|A_1(t)| + \left|\frac{A_2(t)}{A_1(t)}\right|\right). \end{aligned} \tag{5.5}$$

We know, by (5.2) and (5.3), that, for $|t| \le M_1^{-1}|a_1|$, we have $|A_1(t)| \le 2|a_1|$ and $\left|\frac{A_2(t)}{A_1(t)}\right| \le 4|a_1|$. Hence, using this to estimate the right-hand side of (5.5), we get

$$|A_3(t)| \le M_0|a_1|^3,$$

for $|t| \le M_1^{-1}|a_1|$. Now, for $|t| \le M_1^{-1}|a_1|$, consider

$$\begin{aligned}\left|a_3 + a_3^{(1)}t + \frac{a_3^{(2)}}{2!}t^2\right| &= \left|A_3(t) - \frac{A_3^{(3)}(\xi)}{3!}t^3\right| \le |A_3(t)| + \frac{|A_3^{(3)}(\xi)|}{3!}|t|^3 \\ &\le M_0|a_1|^3 + M_0M_1^{-3}|a_1|^3 \le M_0^2|a_1|^3.\end{aligned}$$

Use lemma 5.4.2 to get the following estimates

$$|a_3^{(j)}| \le M_0M_1^{2+j}|a_1|^{3-j} \qquad (0 \le j \le 2). \tag{5.6}$$

Therefore,

$$|A_3''(t)| = |a_3^{(2)} + A_3^{(3)}(\xi)t| \le M_0M_1^4|a_1| + M_0M_1^{-1}|a_1| \le M_0^2M_1^4|a_1|, \tag{5.7}$$

for $|t| \le M_1^{-1}|a_1|$.

Once more let us consider the equation $\frac{\partial}{\partial x}P(t)(x) = 3x^2 + 2A_1(t)x + A_2(t) = 0$ with roots $x_1(t)$ and $x_2(t)$. For the following consideration let us assume that not both roots vanish. We claim that one of the roots has the form $-q\frac{A_2(t)}{A_1(t)}$ with $0 < q \le 1$.

Let t be fixed. If one root vanishes then the statement is trivial. So assume $x_1(t)$ and $x_2(t)$ do not vanish, and without loss of generality let $-\frac{A_2(t)}{A_1(t)} > 0$. Indirectly we suppose no root lies in $[0, -\frac{A_2(t)}{A_1(t)}]$. Then there is a root larger than $-\frac{A_2(t)}{A_1(t)} = \frac{2x_1(t)x_2(t)}{x_1(t)+x_2(t)}$ (otherwise $-\frac{A_2(t)}{A_1(t)} < 0$). If this holds for both, $x_1(t)$ and $x_2(t)$, then we find $x_1(t) > x_2(t)$ and $x_1(t) < x_2(t)$ simultaneously, a contradiction. If, say, $x_1(t) > -\frac{A_2(t)}{A_1(t)}$ and $x_2(t) < 0$, then $2 < \frac{x_1^2(t)+x_1(t)x_2(t)}{x_1(t)x_2(t)} = \frac{x_1(t)}{x_2(t)} + 1 < 1$. This yields the assertion.

Let $x_0(t)$ be a root of $\frac{\partial}{\partial x}P(t)(x) = 0$ of the form $-q\frac{A_2(t)}{A_1(t)}$ $(0 < q \le 1)$ with minimal absolute-value. If $x_0(t) \ne 0$, then $\frac{\partial}{\partial x}P(t)(x)$ has constant sign on the open segment between 0 and $x_0(t)$. Therefore, $0 \ge \frac{\partial^2}{\partial x^2}P(t)(x_0(t)) \cdot x_0(t) \cdot \frac{\partial}{\partial x}P(t)(0) = \frac{\partial^2}{\partial x^2}P(t)(x_0(t)) \cdot \left(-q\frac{A_2(t)}{A_1(t)}\right) \cdot A_2(t)$, implying $0 \le \frac{\partial^2}{\partial x^2}P(t)(x_0(t)) \cdot A_1(t)$. If $x_0(t) = 0$ we come to the same result, since then $\frac{\partial^2}{\partial x^2}P(t)(x_0(t)) \cdot A_1(t) = 2A_1^2(t) \ge 0$.

We want to use this to find an estimate for $A_3(t)A_1(t)$. We may suppose $A_1(t) \ne 0$. Consider the following cases: If $x_0(t) = -q\frac{A_2(t)}{A_1(t)}$ is a root of $P(t)$, then

$$\begin{aligned}|A_3(t)| &\le |x_0(t)|^3 + |A_1(t)||x_0(t)|^2 + |A_2(t)||x_0(t)| \\ &\le \left|\frac{A_2(t)}{A_1(t)}\right|^3 + |A_1(t)|\left|\frac{A_2(t)}{A_1(t)}\right|^2 + |A_2(t)|\left|\frac{A_2(t)}{A_1(t)}\right|.\end{aligned}$$

If $x_0(t)$ is not a root of $P(t)$, then $P(t)(x_0(t)) \ne 0$ and $\frac{\partial^2}{\partial x^2}P(t)(x_0(t)) \ne 0$ must have different signs (see lemma 5.4.4). Assume that $\frac{\partial^2}{\partial x^2}P(t)(x_0(t)) > 0$. Then, by the above, we have $A_1(t) > 0$ and $P(t)(x_0(t)) = x_0^3(t)+A_1(t)x_0^2(t)+A_2(t)x_0(t)+A_3(t) < 0$. Therefore

$$\begin{aligned}
A_3(t)A_1(t) &< \left(-x_0^3(t) - A_1(t)x_0^2(t) - A_2(t)x_0(t)\right) A_1(t) \\
&\leq \left(|x_0(t)|^3 + |A_1(t)||x_0(t)|^2 + |A_2(t)||x_0(t)|\right) |A_1(t)| \\
&\leq \left(\left|\frac{A_2(t)}{A_1(t)}\right|^3 + |A_1(t)| \left|\frac{A_2(t)}{A_1(t)}\right|^2 + |A_2(t)| \left|\frac{A_2(t)}{A_1(t)}\right|\right) |A_1(t)|.
\end{aligned}$$

In an analogous way we get the same estimate, if $\frac{\partial^2}{\partial x^2} P(t)(x_0(t)) < 0$.

At the beginning of these considerations we have excluded the case that both roots of $\frac{\partial}{\partial x} P(t)(x)$ vanish. If both of them vanish, then 0 is a root of $P(t)$, and so $A_3(t) = 0$, whence the above estimate of $A_3(t)A_1(t)$ is trivially fulfilled.

By (5.2), (5.3), and (5.4) together with this estimate, we have, for $|t| \leq M_1^{-1}|a_1|$:

$$\begin{aligned}
A_3(t)a_1 &\leq 4|a_1| \frac{|A_2(t)|^2}{|A_1(t)|} \left(\frac{|A_2(t)|}{|A_1(t)|^2} + 2\right) \\
&\leq 8|A_2(t)|^2 \left(2\left|\frac{a_1}{A_1(t)}\right|^2 + 2\right) \\
&\leq 80|A_2(t)|^2 \leq 320a_2^2.
\end{aligned}$$

We plug $t = \pm M_1^{-3}\left|\frac{a_2}{a_1}\right|$ into this inequality (remember $a_3 = 0$):

$$\pm a_1 a_3^{(1)} M_1^{-3} \left|\frac{a_2}{a_1}\right| + a_1 \frac{A_3^{(2)}(\xi)}{2!} M_1^{-6} \left|\frac{a_2}{a_1}\right|^2 \leq 320a_2^2$$

and obtain, with (5.7),

$$\begin{aligned}
\pm a_1 a_3^{(1)} M_1^{-3} \left|\frac{a_2}{a_1}\right| &\leq 320a_2^2 - a_1 \frac{A_3^{(2)}(\xi)}{2!} M_1^{-6} \left|\frac{a_2}{a_1}\right|^2 \\
&\leq 320a_2^2 + |a_1| \frac{|A_3^{(2)}(\xi)|}{2!} M_1^{-6} \left|\frac{a_2}{a_1}\right|^2 \\
&\leq 320a_2^2 + \frac{1}{2} M_0^2 M_1^{-2} a_2^2 \\
&\leq M_0 a_2^2.
\end{aligned}$$

Hence, $|a_3^{(1)}| \leq M_0 M_1^3 |a_2|$ which shows the statement in case (A).

In case (B), where $|a_2| \geq a_1^2$, we put $t = \pm M_0^{-1}|a_2|^{\frac{1}{2}}$ into (5.1):

$$a_2 \pm a_2^{(1)} M_0^{-1} |a_2|^{\frac{1}{2}} + \frac{A_2^{(2)}(\xi)}{2!} M_0^{-2}|a_2| \leq \frac{1}{3}(a_1 \pm A_1^{(1)}(\eta) M_0^{-1}|a_2|^{\frac{1}{2}})^2.$$

Thus,

$$\begin{aligned}
\pm a_2^{(1)} M_0^{-1}|a_2|^{\frac{1}{2}} &\leq \frac{1}{3}a_1^2 \pm \frac{2}{3} a_1 A_1^{(1)}(\eta) M_0^{-1}|a_2|^{\frac{1}{2}} + \frac{1}{3}(A_1^{(1)}(\eta))^2 M_0^{-2}|a_2| \\
&\quad - a_2 - \frac{A_2^{(2)}(\xi)}{2!} M_0^{-2}|a_2| \\
&\leq \frac{1}{3}a_1^2 + \frac{2}{3}|a_1||A_1^{(1)}(\eta)| M_0^{-1}|a_2|^{\frac{1}{2}} + \frac{1}{3}|A_1^{(1)}(\eta)|^2 M_0^{-2}|a_2| \\
&\quad + |a_2| + \frac{|A_2^{(2)}(\xi)|}{2!} M_0^{-2}|a_2| \\
&\leq M_0|a_2|,
\end{aligned}$$

whence $|a_2^{(1)}| \le M_0^2|a_2|^{\frac{1}{2}}$, which we use to get the following inequalities, for $|t| \le M_2^{-1}|a_2|^{\frac{1}{2}}$:

$$\frac{1}{2} \le 1 - \frac{|a_2^{(1)}|}{|a_2|}|t| - \frac{|A_2^{(2)}(\xi)|}{2|a_2|}|t|^2 \le \frac{A_2(t)}{a_2} \le 1 + \frac{|a_2^{(1)}|}{|a_2|}|t| + \frac{|A_2^{(2)}(\xi)|}{2|a_2|}|t|^2 \le 2, \quad (5.8)$$

since $\frac{|a_2^{(1)}|}{|a_2|}|t| + \frac{|A_2^{(2)}(\xi)|}{2|a_2|}|t|^2 \le M_0^2M_2^{-1} + M_0M_2^{-2} \le \frac{1}{2}$.

Consider, for $|t| \le M_2^{-1}|a_2|^{\frac{1}{2}}$,

$$|A_1(t)| = |a_1 + A_1^{(1)}(\xi)t| \le |a_1| + |A_1^{(1)}(\xi)||t| \le |a_2|^{\frac{1}{2}} + M_0M_2^{-1}|a_2|^{\frac{1}{2}} \le 2|a_2|^{\frac{1}{2}},$$

and, with (5.1),

$$\left|\frac{A_2(t)}{A_1(t)}\right| \le \frac{1}{3}|A_1(t)|.$$

Apply these estimates to (5.5):

$$|A_3(t)| \le M_0|a_2|^{\frac{3}{2}},$$

for $|t| \le M_2^{-1}|a_2|^{\frac{1}{2}}$. Using this, we see that

$$\begin{aligned}\left|a_3 + a_3^{(1)}t + \frac{a_3^{(2)}}{2!}t^2\right| &= \left|A_3(t) - \frac{A_3^{(3)}(\xi)}{3!}t^3\right| \le |A_3(t)| + \frac{|A_3^{(3)}(\xi)|}{3!}|t|^3 \\ &\le M_0|a_2|^{\frac{3}{2}} + M_0M_2^{-3}|a_2|^{\frac{3}{2}} \le M_0^2|a_2|^{\frac{3}{2}},\end{aligned}$$

for $|t| \le M_2^{-1}|a_2|^{\frac{1}{2}}$. By lemma 5.4.2, we get

$$|a_3^{(j)}| \le M_0M_2^{2+j}|a_2|^{\frac{3-j}{2}} \qquad (0 \le j \le 2) \quad (5.9)$$

which concludes case (B).

5.1.2. Note that, if in the assumptions of 5.1.1 we simply replace $A_2(0) \ne 0$ by $A_1(0) \ne 0$ and $A_2(0) = 0$, then there exists a positive constant C such that

$$\left|\frac{A_3''(0)}{A_1(0)}\right| \le \left(\max_{i,t,j\le i}|A_i^{(j)}(t)| + 2\right)^C.$$

In the case (A) this corresponds to (5.6). Case (B) does not appear, since $a_2 = 0$ would imply $a_1 = 0$, contrary to the assumption.

5.1.3. We will need similar results to those in 5.1.1 and 5.1.2 also for the degrees one and two. But these are more easy to get: For $P(t)(x) = x + A_1(t)$, where $A_1 \in C^1([-1,1])$, we have, of course, $|A_1'(0)| \le \max_{j=0,1;t}|A_1^{(j)}(t)|$.

The roots of $P(t)(x) = x^2 + A_1(t)x + A_2(t)$ are always real, if and only if $A_1^2(t) - 4A_2(t) \ge 0$. Moreover, suppose that $A_i \in C^i([-1,1])$, for $i = 1,2$, $A_1(0) \ne 0$ and $A_2(0) = 0$. Put $M_0 = 4(\max_{i,t,j\le i}|A_i^{(j)}(t)| + 4)^4$ here. Set $t = \pm M_0^{-1}a_1$ in the previous inequality:

$$\pm a_2^{(1)}M_0^{-1}a_1 + \frac{A_2^{(2)}(\xi)}{2!}M_0^{-2}a_1^2 \le \frac{1}{4}\left(a_1 \pm A_1^{(1)}(\eta)M_0^{-1}a_1\right)^2.$$

Therefore,

$$\begin{aligned}
\pm a_2^{(1)} M_0^{-1} a_1 &\le \frac{1}{4}a_1^2 \pm \frac{1}{2}A_1^{(1)}(\eta)M_0^{-1}a_1^2 + \frac{1}{4}\left(A_1^{(1)}(\eta)\right)^2 M_0^{-2}a_1^2 \\
&\quad - \frac{A_2^{(2)}(\xi)}{2!}M_0^{-2}a_1^2 \\
&\le \frac{1}{4}a_1^2 + \frac{1}{2}|A_1^{(1)}(\eta)|M_0^{-1}a_1^2 + \frac{1}{4}|A_1^{(1)}(\eta)|^2 M_0^{-2}a_1^2 \\
&\quad + \frac{|A_2^{(2)}(\xi)|}{2!}M_0^{-2}a_1^2 \\
&\le \frac{1}{4}a_1^2 + \frac{1}{4}a_1^2 + \frac{1}{2}a_1^2 + M_0^{-1}a_1^2 \\
&\le 2a_1^2,
\end{aligned}$$

whence

$$|a_2^{(1)}| \le 2M_0|a_1|.$$

Estimates of the kind as in 5.1.2 are trivial for degrees one and two.

5.1.4. Suppose all roots of $P(t)(x) = x^3 - a_1(t)x^2 + a_2(t)x - a_3(t)$ are real for each $t \in (-1,1)$, and $a_i \in C^3((-1,1))$, for $1 \le i \le 3$. *We assert that for any compact subset $K \subseteq (-1,1)$ there exists a constant C_K such that all (differentiably chosen) roots x_j, $1 \le j \le 3$, of P satisfy $\left|x_j'(t)\right| < C_K$ for all $t \in K$.*

For contradiction suppose $x_j'(t)$ is unbounded on a compact subset $K \subseteq (-1,1)$ for some $1 \le j \le 3$. Without loss of generality, say $j = 1$, and assume for contradiction that there is a sequence $(t_p)_{p\in\mathbb{N}}$ in K such that $t_p \to t_\infty$, $x_1(t_p) \to x_1(t_\infty)$, and $|x_1'(t_p)| \to \infty$, as $p \to \infty$. By switching to a subsequence, we can achieve that $x_1(t_p)$ has fixed multiplicity q for all $p \in \mathbb{N}$, and $x_1(t_\infty)$ has multiplicity $s \ge q$. Consider

$$\begin{aligned}
Q_p(t)(\tilde{x}) &= P(t)(\tilde{x} + x_1(t_p)) \\
&= \tilde{x}^3 + \underbrace{\frac{1}{2!}\frac{\partial^2}{\partial x^2}P(t)(x_1(t_p))}_{=b_{p,1}(t)}\,\tilde{x}^2 + \underbrace{\frac{\partial}{\partial x}P(t)(x_1(t_p))}_{=b_{p,2}(t)}\,\tilde{x} + \underbrace{P(t)(x_1(t_p))}_{=b_{p,3}(t)}.
\end{aligned}$$

Moreover, we define $b_{p,0} = 1$. As we will see in theorem 5.2.1, $x_1'(t_p)$ has to satisfy the following equation:

$$T_p(x) = b_{p,3-q}(t_p)x^q + \frac{1}{1!}b_{p,3-q+1}^{(1)}(t_p)x^{q-1} + \cdots + \frac{1}{q!}b_{p,3}^{(q)}(t_p) = 0 \qquad (p \in \mathbb{N}).$$

Our goal is to estimate the coefficients of $b_{p,3-q}(t_p)^{-1}T_p(x)$. If we can show that they are bounded, then also $x_1'(t_p)$ is bounded (see lemma 5.4.3), and we are done.

Observe that

$$b_{p,3-q}(t_p) = \frac{1}{q!}\frac{\partial^q}{\partial x^q}P(t_p)(x_1(t_p)) \ne 0,$$

and

$$b_{p,3-q+j}(t_p) = \frac{1}{(q-j)!}\frac{\partial^{q-j}}{\partial x^{q-j}}P(t_p)(x_1(t_p)) = 0 \qquad (0 < j \le q),$$

since $x_1(t_p)$ has multiplicity q. We differentiate $Q_p(t)(\tilde{x})$:

$$\left(\frac{\partial}{\partial \tilde{x}}\right)^{q-j} Q_p(t)(\tilde{x}) = \frac{3!}{(3-q+j)!}\tilde{x}^{3-q+j} + \cdots + (q-j)!b_{p,3-q+j}(t),$$

where $j = 1, 2$ and $q \ge j$. Note that this hyperbolic polynomial has at most degree three, and all coefficients are of class C^3. Let us apply the results of 5.1.1 and 5.1.2

to it, for $j = 1$ and $j = 2$, respectively. We find that

$$\left| \frac{b^{(j)}_{p,3-q+j}(t_p)}{b_{p,3-q}(t_p)} \right| \le C \qquad (j = 1, 2),$$

where C does not depend on p, since $t_p \in K$ which is compact. If $q < 3$ we are done. If $q = 3$, then

$$\left| \frac{b^{(3)}_{p,3}(t_p)}{b_{p,0}(t_p)} \right| = |b^{(3)}_{p,3}(t_p)|$$

is also bounded, since $b^{(3)}_{p,3}$ is continuous and $t_p \in K$. This shows the assertion.

5.2. Differentiability of the roots

In this section we give Bronshtein's proof of the fact that the roots of a C^n curve of hyperbolic polynomials of degree n may be chosen differentiable, and that the potential derivatives of a root must satisfy a polynomial equation, namely (5.10). Compare with theorem 4.3.2.

5.2.1. Theorem. *Suppose that for any $t \in (-1, 1)$ the polynomial*

$$P(t)(x) = x^n - a_1(t)x^{n-1} + \cdots + (-1)^n a_n(t)$$

is hyperbolic and the multiplicities of its roots do not exceed k. We assume that the coefficients a_i are of class C^k on $(-1, 1)$, for $1 \le i \le n$. Then for any point $t_0 \in (-1, 1)$ there exist n roots $x_j = x_j(t_0; t)$, $1 \le j \le n$, of $P(t)$ which are differentiable at $t = t_0$.

Moreover, each of the q possible derivatives at t_0 of a q-fold root $x(t_0)$ of $P(t_0)$ satisfies the following hyperbolic equation:

$$b^{(0)}_0(t_0)x^q + \frac{1}{1!}b^{(1)}_1(t_0)x^{q-1} + \cdots + \frac{1}{q!}b^{(q)}_q(t_0) = 0, \tag{5.10}$$

where

$$b_i(t) = \frac{1}{(q-i)!}\left(\frac{\partial}{\partial x}\right)^{q-i}\Bigg|_{x=x(t_0)} P(t)(x) \qquad (0 \le i \le q).$$

Proof. Without loss of generality we can assume that $t_0 = 0$. Let x_0 be a root of the polynomial P(0) of multiplicity q. Consider the following, still hyperbolic, polynomial

$$Q(t)(x) = P(t)(x + x_0) = \sum_{j=0}^{n-q-1} \tilde{a}_j(t)x^{n-j} + \sum_{j=0}^{q} b_j(t)x^{q-j}. \tag{5.11}$$

Note that $\tilde{a}_0 = 1$. The coefficients $\tilde{a}_j$ and b_j are again of class C^k on $(-1, 1)$ and, by Taylor's formula,

$$b_j(t) = \frac{1}{(q-j)!}\left(\frac{\partial}{\partial x}\right)^{q-j}\Bigg|_{x=x_0} P(t)(x) \qquad (j = 0, \ldots, q). \tag{5.12}$$

Of course, the $\tilde{a}_j$ are Taylor-coefficients as well, but we are not interested in their explicit form. Put

$$b^{(i)}_j = \left(\frac{d}{dt}\right)^{i}\Bigg|_{t=0} b_j(t) \qquad (0 \le i, j \le q).$$

Since x_0 is a q-fold root of $P(0)$, we find $b^{(0)}_0 \neq 0$ and $b^{(0)}_j = 0$, for $1 \le j \le q$.

Claim (1). *For $1 \le j \le q$ and $0 \le i \le j - 1$, we have $b^{(i)}_j = 0$.*

This assertion is equivalent to the statement that any b_j can be presented near 0 in the form $b_j(t) = t^j \tilde{b}_j(t)$, where $\tilde{b}_j$ is a continuous function. Assume the assertion is wrong. Let j_0 be the minimal index in $\{1, \ldots, q\}$ for which claim 1 is not true. Thus, there exists an $0 \le i_0 \le j_0 - 1$ such that $b_{j_0}^{(i)} = 0$ for each $i < i_0$ and $b_{j_0}^{(i_0)} \ne 0$.

Consider the polynomial $|t|^{-i_0}(\frac{\partial}{\partial x})^{q-j_0} Q(t)(x)$ and replace x by $|t|^{\frac{i_0}{j_0}} \tilde{x}$. By lemma 5.4.4(1), the resulting polynomial $R(t)(\tilde{x})$ is hyperbolic, and it takes the following form

$$\begin{aligned} R(t)(\tilde{x}) = &\sum_{j=0}^{n-q-1} \frac{(n-j)!}{(n-q-j+j_0)!} \tilde{a}_j(t)|t|^{\frac{i_0}{j_0}(n-q-j)} \tilde{x}^{n-q-j+j_0} \\ &+ \sum_{j=0}^{j_0} \frac{(q-j)!}{(j_0-j)!} b_j(t)|t|^{-\frac{i_0}{j_0}j} \tilde{x}^{j_0-j}. \end{aligned} \tag{5.13}$$

We analyze the coefficients in the second sum of (5.13):

For $j = 0$,

$$b_0(t)|t|^0 = b_0(t) = b_0^{(0)} + b_0^{(1)}(\xi)t,$$

where the second term is continuous in t and vanishes for $t = 0$.

For $0 < j < j_0$, we find, by assumption,

$$\begin{aligned} b_j(t)|t|^{-\frac{i_0}{j_0}j} &= \left(b_j^{(0)} + \cdots + \frac{b_j^{(j-1)}}{(j-1)!} t^{j-1} + \frac{b_j^{(j)}(\xi)}{j!} t^j \right) |t|^{-\frac{i_0}{j_0}j} \\ &= \frac{1}{j!} b_j^{(j)}(\xi) \cdot \operatorname{sgn}\left(t^{-\frac{i_0}{j_0}j}\right) \cdot t^{j-\frac{i_0}{j_0}j} \end{aligned}$$

which is continuous in t and vanishes for $t = 0$.

For $j = j_0$, we get

$$\begin{aligned} b_{j_0}(t)|t|^{-i_0} &= \left(b_{j_0}^{(0)} + \cdots + \frac{b_{j_0}^{(i_0)}}{i_0!} t^{i_0} + \frac{b_{j_0}^{i_0+1}(\xi)}{(i_0+1)!} t^{i_0+1} \right) |t|^{-i_0} \\ &= \frac{b_{j_0}^{(i_0)}}{i_0!} \cdot \operatorname{sgn}\left(t^{-i_0}\right) + \frac{b_{j_0}^{i_0+1}(\xi)}{(i_0+1)!} \cdot \operatorname{sgn}\left(t^{-i_0}\right) \cdot t, \end{aligned}$$

where the second term is again continuous in t and vanishes for $t = 0$.

Clearly, the coefficients in the first sum of (5.13) are continuous in t and vanish for $t = 0$, too. Thus, (5.13) can be written as follows:

$$R(t)(\tilde{x}) = \frac{q!}{j_0!} b_0^{(0)} \tilde{x}^{j_0} + \frac{(q-j_0)!}{i_0!} b_{j_0}^{(i_0)} \operatorname{sgn}\left(t^{i_0}\right) + c_0(t)\tilde{x}^{n-q+j_0} + \cdots + c_{n-q+j_0}(t), \tag{5.14}$$

where $c_0, \ldots, c_{n-q+j_0}$ are continuous functions in t, and all of them vanish for $t = 0$.

Before we finish the proof of claim 1, we have to consider the following assertion:

Claim (2). *The equation* $b_0^{(0)} \tilde{x}^{j_0} + b_{j_0}^{(i_0)} \operatorname{sgn}\left(t^{i_0}\right) = 0$, *where* $0 < i_0 < j_0 \ge 2$, *has non-real roots, for* $tb_0^{(0)} b_{j_0}^{(i_0)} > 0$.

If j_0 is odd, then $j_0 \ge 3$, and the equation has non-real roots whenever $t \ne 0$, in particular, when $tb_0^{(0)} b_{j_0}^{(i_0)} > 0$. If j_0 is even, let us first consider the case $j_0 = 2$. But then $i_0 = 1$, and so there exist non-real roots, if $\operatorname{sgn}(t)\, b_0^{(0)} b_{j_0}^{(i_0)} > 0$ which is equivalent to $tb_0^{(0)} b_{j_0}^{(i_0)} > 0$. The case where j_0 is even and $j_0 \ge 4$ can be reduced to the considered cases, by substitution. Thus, claim 2 is proved.

Now, consider the polynomial $R(t)(\tilde{x})$ in (5.14). For t near 0 such that the condition $tb_0^{(0)}b_{j_0}^{(i_0)} > 0$ in claim 2 is satisfied, theorem 2.2.3 implies that the polynomial $R(t)(\tilde{x})$ has non-real roots, a contradiction. Thus, claim 1 follows.

Putting $x = t\tilde{x}$ in equation (5.11) and dividing it by t^q, we obtain:

$$t^{-q}Q(t)(t\tilde{x}) = \sum_{j=0}^{n-q-1} \tilde{a}_j(t)t^{n-q-j}\tilde{x}^{n-j} + \sum_{j=0}^{q} b_j(t)t^{-j}\tilde{x}^{q-j}.$$

Applying claim 1, we get, for all $0 \le j \le q$,

$$b_j(t)t^{-j} = \left(b_j^{(0)} + \cdots + \frac{b_j^{(j)}}{j!}t^j + o(t^j)\right)t^{-j} = \frac{b_j^{(j)}}{j!} + t^{-j}o(t^j),$$

where the second term is continuous in t and vanishes for t approaching 0.

This implies that

$$t^{-q}Q(t)(t\tilde{x}) = b_0^{(0)}\tilde{x}^q + \frac{1}{1!}b_1^{(1)}\tilde{x}^{q-1} + \cdots + \frac{1}{q!}b_q^{(q)} + d_0(t)\tilde{x}^n + \cdots + d_n(t),$$

where $d_0, \ldots, d_n$ are continuous functions in t, and $d_j(0) = 0$, for all $0 \le j \le n$. Using theorem 2.2.3 again we find that the polynomial $t^{-q}Q(t)(t\tilde{x})$, in a sufficiently small neighborhood of $t = 0$, has q (with multiplicities) roots $\tilde{x}_1(t), \ldots, \tilde{x}_q(t)$ which are continuous at $t = 0$. All of them are real, since $t^{-q}Q(t)(t\tilde{x})$ is hyperbolic, by construction. Then

$$0 = t^{-q}Q(t)(t\tilde{x}_j(t)) = t^{-q}P(t)(x_0 + t\tilde{x}_j(t)) \qquad (1 \le j \le q)$$

implies that, for t near 0, $P(t)$ has q roots of the form $x_j(t) = x_0 + t\tilde{x}_j(t)$, where $1 \le j \le q$. They coincide for $t = 0$ and are differentiable at this point,

$$\lim_{t\to 0} \frac{x_j(t) - x_j(0)}{t} = \lim_{t\to 0} \frac{t\tilde{x}_j(t)}{t} = \tilde{x}_j(0),$$

with derivative $\tilde{x}_j(0)$ which satisfies the following equation:

$$b_0^{(0)}x^q + \frac{1}{1!}b_1^{(1)}x^{q-1} + \cdots + \frac{1}{q!}b_q^{(q)} = 0,$$

with

$$b_i^{(i)} = \frac{1}{(q-i)!}\left(\frac{\partial}{\partial t}\right)^i\bigg|_{t=0}\left(\frac{\partial}{\partial x}\right)^{q-i}\bigg|_{x=x_0} P(t)(x) \qquad (0 \le i \le q).$$

Therefore, the theorem is proved. □

5.3. A comparison

Let us compare here the methods, Alekseevsky, Kriegl, Losik, and Michor use to show that the roots of a C^n curve of hyperbolic polynomials of degree n may be chosen differentiably on the one hand (theorem 4.3.2), with those Bronshtein uses on the other hand (theorem 5.2.1). It will turn out that the two approaches are very similar with the decisive difference that the iterative method in the proof of theorem 4.3.2, following the algorithm 4.2.9, is done simultaneously by Bronshtein.

In the following we shall repeat the main steps in Bronshtein's proof of theorem 5.2.1 and comment them from Alekseevsky, Kriegl, Losik, and Michor's point of view. The main ingredients of their proof of theorem 4.3.2 are the splitting lemma 4.2.3, the multiplicity lemma 4.2.8, and theorem 2.3.2 which provides a continuous parameterization of the roots.

The structure of Bronshtein's proof is the following: He fixes a q-fold root x_0 of $P(0)$, and he is going to show that $P(t)$ has q roots $x_1(t), \ldots, x_q(t)$ for t near 0 which agree for $t = 0$ and are differentiable there. Here implicitly is used the splitting lemma 4.2.3. But note that, differently from the use Alekseevsky, Kriegl, Losik, and Michor make of it, in the following steps the curve $P(t)$ will not be factorized.

Next he puts

$$Q(t)(x) = P(t)(x + x_0) = \tilde{a}_0(t)x^n + \cdots + \tilde{a}_{n-q-1}(t)x^{q+1} + b_0(t)x^q + \cdots + b_q(t),$$

where

$$b_j(t) = \frac{1}{(q-j)!}\left(\frac{\partial}{\partial x}\right)^{q-j}\Bigg|_{x=x_0} P(t)(x) \qquad (0, \le j \le q),$$

and so he gains that then $x = 0$ is a q-fold root of $Q(0)$. This shifting of the focal point to 0 is closely related to the change of variables $x \rightsquigarrow x + \frac{a_1(t)}{n}$, equivalently, the assumption $a_1 = 0$ in the proof of theorem 4.3.2.

Claim 1 states that $b_j^{(i)}(0) = 0$, for $1 \le j \le q$ and $0 \le i \le j-1$, which is equivalent to the statement that each b_j can be presented near 0 in the form $b_j(t) = t^j \tilde{b}_j(t)$ for a continuous function $\tilde{b}_j$. Hence claim 1 corresponds to the multiplicity lemma 4.2.8.

The next important step in Bronshtein's proof is to consider $t^{-q}Q(t)(t\tilde{x})$ which, with claim 1, takes the following form

$$t^{-q}Q(t)(t\tilde{x}) = b_0^{(0)}(0)\tilde{x}^q + \frac{1}{1!}b_1^{(1)}(0)\tilde{x}^{q-1} + \cdots + \frac{1}{q!}b_q^{(q)}(0) + d_0(t)\tilde{x}^n + \cdots + d_n(t),$$

where $d_0, \dots, d_n$ are continuous functions in t, and $d_j(0) = 0$, for all $0 \le j \le n$. This continuous curve of hyperbolic polynomials $t^{-q}Q(t)(t\tilde{x})$ corresponds to $P_1(t)(x)$ in the proof of theorem 4.3.2. The intended purpose of $t^{-q}Q(t)(t\tilde{x})$ and $P_1(t)(x)$ in the respective proofs is the same: Their roots $\tilde{x}_1(t), \dots, \tilde{x}_q(t)$ may be chosen continuous, by theorem 2.2.3 and theorem 2.3.2, respectively, such that $x_j(t) = x_0 + t\tilde{x}_j(t)$, $1 \le j \le q$, are q roots of $P(t)$, for t near 0, which are differentiable at $t = 0$ and coincide at $t = 0$.

Moreover, Bronshtein can conclude that the q possible derivatives at $t = 0$ of the q-fold root x_0 of $P(0)$ satisfy the following hyperbolic equation:

$$b_0^{(0)}(0)x^q + \frac{1}{1!}b_1^{(1)}(0)x^{q-1} + \cdots + \frac{1}{q!}b_q^{(q)}(0) = 0.$$

The crucial point here is that this property is valid for any q-fold root of $P(t_0)$, where t_0 is arbitrary; then its possible derivatives have to fulfill

$$b_0^{(0)}(t_0)x^q + \frac{1}{1!}b_1^{(1)}(t_0)x^{q-1} + \cdots + \frac{1}{q!}b_q^{(q)}(t_0) = 0.$$

So this equation accounts for its dependence on the parameter t. This will be of decisive importance in the proof of theorem 5.5.13, when we deal with the local boundedness of the derivatives of the roots.

In the approach of Alekseevsky, Kriegl, Losik and Michor we have a similar statement: We know that the roots of $P_1(0)(x) = 0$ are the possible derivatives of the only root 0 of $P(0)$ (remember that here we are in the case $a_2(0) = 0$ and we probably have already used the splitting lemma 4.2.3 such this P is not the curve of polynomials we have started from). Since we have applied the splitting lemma 4.2.3 before, $P_1(t)$ is defined only on a small open interval, but in view of the local boundedness of the derivatives of the roots we would need a statement for the whole domain of the parameter t.

5.4. The interrelation of coefficients and roots of hyperbolic polynomials

In this section are collected the preliminaries used in section 5.5.

Recall that any monic polynomial P over $\mathbb{C}$ of degree n with roots $x_1, \dots, x_n$ can be presented as

$$P(x) = x^n - a_1 x^{n-1} + \cdots + (-1)^n a_n = \prod_{i=1}^{n} (x - x_i).$$

By carrying out the multiplications on the right-hand side and equating coefficients, we find Vieta's formulas

$$a_i = \sum_{1 \le j_1 < \cdots < j_i \le n} x_{j_1} \cdots x_{j_i} \qquad (1 \le i \le n).$$

So we see that the coefficients of P are (up to their sign) the elementary symmetric functions in its roots.

5.4.1. Lemma. *Let the roots x_i of the polynomial*

$$P(x) = x^n - a_1 x^{n-1} + \cdots + (-1)^n a_n = \prod_{i=1}^{n} (x - x_i) \qquad (a_i, x_i \in \mathbb{C})$$

satisfy the inequalities $|x_1| \le |x_2| \le \cdots \le |x_n|$. Then we have

$$|x_2| \le 2n^2 \left(\min \left(\left| \frac{a_n}{a_{n-1}} \right|, \left| \frac{a_n}{a_{n-2}} \right|^{\frac{1}{2}} \right) + \min \left(\left| \frac{a_{n-1}}{a_{n-2}} \right|, \left| \frac{a_{n-1}}{a_{n-3}} \right|^{\frac{1}{2}} \right) \right).$$

Proof. First let us assume that a_{n-1}, a_{n-2}, and a_{n-3} do not vanish. With this assumption we consider the following two cases:

(1) If $2n|x_1| \ge |x_2|$: Vieta's formulas imply:

$$|a_n| = |x_1 x_2 \cdots x_n|,$$

$$|a_{n-1}| = \left| \sum_{j_1 < \cdots < j_{n-1}} x_{j_1} \cdots x_{j_{n-1}} \right| \le \sum_{j_1 < \cdots < j_{n-1}} |x_{j_1} \cdots x_{j_{n-1}}| \le n |x_2 \cdots x_n|,$$

and analogously

$$|a_{n-2}| \le \frac{n(n-1)}{2} |x_3 \cdots x_n| \le n^2 |x_3 \cdots x_n|.$$

In particular one sees that, since $a_{n-1} \ne 0$, none of $x_2, \dots, x_n$ vanishes. Then $|x_1| \le n \left| \frac{a_n}{a_{n-1}} \right|$, and $|x_1|^2 \le |x_1 x_2| \le n^2 \left| \frac{a_n}{a_{n-2}} \right|$. This implies

$$|x_2| \le 2n|x_1| \le 2n^2 \min \left(\left| \frac{a_n}{a_{n-1}} \right|, \left| \frac{a_n}{a_{n-2}} \right|^{\frac{1}{2}} \right).$$

(2) If $2n|x_1| < |x_2|$: With Vieta's formulas we find:

$$|a_{n-1}| = \left| (x_2 \cdots x_n) \left(1 + \frac{x_1}{x_2} + \cdots + \frac{x_1}{x_n} \right) \right|,$$

$$|a_{n-2}| \le \frac{n(n-1)}{2} |x_3 \cdots x_n|,$$

$$|a_{n-3}| \le \frac{n(n-1)(n-2)}{6} |x_4 \cdots x_n|.$$

We have

$$\left| 1 + \frac{x_1}{x_2} + \cdots + \frac{x_1}{x_n} \right| \ge \left| 1 - \left| \frac{x_1}{x_2} + \cdots + \frac{x_1}{x_n} \right| \right| > \frac{1}{2},$$

since

$$\left|\frac{x_1}{x_2}+\cdots+\frac{x_1}{x_n}\right| \le \left|\frac{x_1}{x_2}\right|+\cdots+\left|\frac{x_1}{x_n}\right| < (n-1)\frac{1}{2n} < \frac{1}{2}.$$

Thus,

$$2|a_{n-1}| > |x_2\cdots x_n|.$$

Therefore, we obtain

$$|x_2| = \frac{|x_2\cdots x_n|}{|x_3\cdots x_n|} < \frac{2|a_{n-1}|}{|a_{n-2}|}\cdot\frac{n(n-1)}{2} < n^2\left|\frac{a_{n-1}}{a_{n-2}}\right|$$

and

$$|x_2|^2 \le |x_2x_3| = \frac{|x_2\cdots x_n|}{|x_4\cdots x_n|} < \frac{2|a_{n-1}|}{|a_{n-3}|}\cdot\frac{n(n-1)(n-2)}{6} < n^4\left|\frac{a_{n-1}}{a_{n-3}}\right|.$$

Then

$$|x_2| < n^2\min\left(\left|\frac{a_{n-1}}{a_{n-2}}\right|, \left|\frac{a_{n-1}}{a_{n-3}}\right|^{\frac{1}{2}}\right).$$

Thus, in both cases the statement is proved.

Now we have to discuss the remaining cases:

- $a_{n-1} = a_n = 0$: Then 0 is an at least 2-fold root of P and the statement of the lemma is trivial.
- $a_{n-1} = 0$, $a_n \neq 0$, and $a_{n-2} = 0$: The first minimum is $+\infty$, so the inequality holds true.
- $a_{n-1} = 0$, $a_n \neq 0$, and $a_{n-2} \neq 0$: In this case the first minimum becomes $\left|\frac{a_n}{a_{n-2}}\right|^{\frac{1}{2}}$. If $2n|x_1| \ge |x_2|$, the statement follows by (1). The case $2n|x_1| < |x_2|$ is impossible, since $0 = 2|a_{n-1}| > |x_2\cdots x_n|$ would imply $a_n = 0$.
- $a_{n-2} = a_n = 0$: Then $x_1 = 0$. So, if $2n|x_1| \ge |x_2|$, the statement is trivial. If $2n|x_1| < |x_2|$, repeat case (2): $a_{n-1} = 0$ would imply that 0 is a 3-fold root; for $a_{n-1} \neq 0$, the inequality $|x_2| \le n^2\left|\frac{a_{n-1}}{a_{n-2}}\right|$ is clear and $|x_2| \le n^2\left|\frac{a_{n-1}}{a_{n-3}}\right|^{\frac{1}{2}}$ is either trivial (for $a_{n-3} = 0$) or was derived in (2).
- $a_{n-2} = 0$, $a_n \neq 0$, $a_{n-1} \neq 0$, and $a_{n-3} = 0$: The second minimum is $+\infty$.
- $a_{n-2} = 0$, $a_n \neq 0$, $a_{n-1} \neq 0$, and $a_{n-3} \neq 0$: Just repeat cases (1) and (2).
- $a_{n-3} = 0$, $a_{n-1} \neq 0$, and $a_{n-2} \neq 0$: (1) and (2) imply the statement.

Hence, all cases are discussed. □

5.4.2. Lemma. *Let $P(x) = a_0 + a_1x + \cdots + a_nx^n$ be a polynomial over $\mathbb{R}$, satisfying $|P(x)| \le C$ for all $|x| \le D$, for positive constants C and D. Then*

$$|a_j| \le 8n^{n+1}\frac{C}{D^j} \qquad (0 \le j \le n).$$

Proof. The condition $|P(x)| \le C$, for all $|x| \le D$, is equivalent to $|\frac{1}{C}P(Dy)| \le 1$, for all $|y| \le 1$.

We recall a result on extremal properties of *Chebyshev polynomials*, see e.g. [**Riv74**]: *Let $\mathcal{P}_n$ be the set of polynomials with maximal degree n. For the Chebyshev polynomial of degree n*

$$T_n(x) = \cos n\theta = t_0^{(n)} + t_1^{(n)}x + \cdots + t_n^{(n)}x^n \qquad (x = \cos\theta)$$

we have

$$t_{n-(2k+1)}^{(n)} = 0 \quad \textit{for} \quad 0 \le k \le \left[\frac{n-1}{2}\right]$$

and

$$t_{n-(2k)}^{(n)} = (-1)^k \sum_{j=k}^{[\frac{n}{2}]} \binom{n}{2j}\binom{j}{k} \quad \textit{for} \quad 0 \le k \le \left[\frac{n}{2}\right].$$

The extrema of $T_n(x)$ *are given by* $\eta_j^{(n)} = \cos\frac{j\pi}{n}$, *where* $0 \le j \le n$. *All of them lie in the interval* $[-1,1]$.
Let $\mathcal{C}_n = \{P \in \mathcal{P}_n : \max_{0\le j\le n} |P(\eta_j^{(n)})| \le 1\}$ *and consider a polynomial* $P(x) = a_0 + a_1 x + \cdots + a_{n+1}x^{n+1}$. *If* $n+1-j$ *is even (or zero) and* $P \in \mathcal{C}_{n+1}$, *then*

$$|a_j| \le \left| t_j^{(n+1)} \right|.$$

If $n+1-j$ *is odd and* $P \in \mathcal{C}_n$, *then*

$$|a_j| \le \left| t_j^{(n)} \right|.$$

By assumption, the polynomial

$$\frac{1}{C}P(Dy) = \frac{a_0}{C} + \frac{a_1 D^1}{C}y + \cdots + \frac{a_n D^n}{C}y^n$$

belongs to $\mathcal{C}_{n+1}$ and $\mathcal{C}_n$. Therefore, and since $\binom{p}{q} \le 2^p$,

$$\left|\frac{a_j D^j}{C}\right| \le \max\left\{\left|t_j^{(n+1)}\right|, \left|t_j^{(n)}\right|\right\} \le n2^{n+1+[\frac{n+1}{2}]} \le 8n^{n+1} \qquad (0 \le j \le n).$$

This completes the proof. □

There is a more elementary proof, too. It does not need those results on Chebyshev polynomials but uses some simple facts from interpolation theory.

Alternative proof. Choose $n+1$ different nodes $-D = x_0 < \cdots < x_n = D$ and consider Newton's form of the interpolating polynomial of degree n

$$N(x) = P(x_0) + P(x_0,x_1)(x-x_0) + \cdots + P(x_0,\ldots,x_n)(x-x_0)\cdots(x-x_{n-1})$$

with the divided differences given by

$$\frac{P(x_{j_0}, x_{j_0+1}, \ldots, x_{j_0+k-1}) - P(x_{j_0+1}, \ldots, x_{j_0+k})}{x_{j_0} - x_{j_0+k}} = P(x_{j_0}, \ldots, x_{j_0+k}),$$

for $0 \le j_0 \le n-1$ and $1 \le k \le n-j_0$. Suppose the nodes are distributed equidistantly. By induction on k we show that

$$|P(x_{j_0}, \ldots, x_{j_0+k})| \le \frac{n^k}{k!}\cdot\frac{C}{D^k} \qquad (0 \le j_0 \le n, 0 \le k \le n-j_0).$$

The case $k=0$ is trivial, since $|P(x_i)| \le C$, for all $0 \le i \le n$, by assumption. Let us assume the statement is true for $k-1$. Then, since the nodes are distributed equidistantly:

$$\begin{aligned}
|P(x_{j_0}, \ldots, x_{j_0+k})| &= \frac{|P(x_{j_0}, \ldots, x_{j_0+k-1}) - P(x_{j_0+1}, \ldots, x_{j_0+k})|}{|x_{j_0} - x_{j_0+k}|} \\
&\le \frac{n}{2kD}\left(|P(x_{j_0}, \ldots, x_{j_0+k-1})| + |P(x_{j_0+1}, \ldots, x_{j_0+k})|\right) \\
&\le \frac{n}{2kD}\cdot 2\cdot\frac{n^{k-1}}{(k-1)!}\cdot\frac{C}{D^{k-1}} = \frac{n^k}{k!}\cdot\frac{C}{D^k}.
\end{aligned}$$

By expanding $N(x)$, we obtain $N(x) = b_0 + b_1 x + \cdots + b_n x^n$ with

$$b_j = \sum_{k=j}^{n} (-1)^{k-j} P(x_0, \ldots, x_k) \sum_{0 \le l_1 < \cdots < l_{k-j} \le k-1} x_{l_1}\cdots x_{l_{k-j}},$$

since the contribution to b_j of each summand of $N(x)$ can be expressed by Vieta's formulas. A polynomial of degree n given by its values at $n+1$ different nodes is unique, thus $P = N$. So we have $a_j = b_j$ for all $0 \le j \le n$, and hence

$$\begin{aligned} |a_j| &\le \sum_{k=j}^{n} |P(x_0, \dots, x_k)| \sum_{0 \le l_1 < \dots < l_{k-j} \le k-1} |x_{l_1}| \cdots |x_{l_{k-j}}| \\ &\le \sum_{k=j}^{n} \frac{n^k}{k!} \cdot \frac{C}{D^k} \cdot \binom{k}{k-j} \cdot D^{k-j} \\ &\le \frac{C}{D^j} \sum_{k=j}^{n} n^k \le n^{n+1} \frac{C}{D^j}. \end{aligned}$$

So the proof is complete. □

5.4.3. Lemma. *For a sequence $(P_m)_{m\in\mathbb{N}}$ of polynomials over $\mathbb{C}$*

$$P_m(x) = x^n + a_{m,1}x^{n-1} + \dots + a_{m,n}$$

with bounded coefficients $a_{m,1}, \dots, a_{m,n}$, the roots $x_{m,1}, \dots, x_{m,n}$ are bounded, too.

Proof. Suppose there is an unbounded sequence $(x_m)_{m\in\mathbb{N}}$ of roots of $(P_m)_m$, i.e.,

$$x_m^n + a_{m,1}x_m^{n-1} + \dots + a_{m,n} = 0 \quad (m \in \mathbb{N}).$$

Consequently,

$$|x_m|^n \le |a_{m,1}||x_m|^{n-1} + \dots + |a_{m,n}| \quad (m \in \mathbb{N}).$$

Without loss of generality we can assume that $(|x_m|)_m$ is strictly increasing and always positive. Thus,

$$|x_m| \le |a_{m,1}| + |a_{m,2}||x_m|^{-1} + \dots + |a_{m,n}||x_m|^{-n+1} \qquad (m \in \mathbb{N}).$$

But the right-hand side is bounded, contradicting the assumption that $(x_m)_m$ is unbounded. □

5.4.4. Lemma. *A hyperbolic polynomial*

$$P(x) = x^n - a_1x^{n-1} + \dots + (-1)^n a_n$$

with real coefficients a_i satisfies the following properties:

(1) *The polynomial $P' = \frac{d}{dx}P$ is hyperbolic, and between any two neighboring roots $x_1 < x_2$ of P there is precisely one (simple) root of P' distinct from x_1 and x_2.*
(2) *Between any two roots $y_1 \le y_2$ (equality means a multiple root) of P' there is a root of P.*
(3) *If $P'(y_0) = 0$ and $P(y_0) \ne 0$, then $P(y_0)P''(y_0) < 0$.*
(4) *If $a_n \ne 0$, then $|a_j| + |a_{j+1}| \ne 0$, for all $1 \le j \le n-1$.*
(5) *If $a_{n-1} \ne 0$, then P has a root of the form $x_0 = n\rho\frac{a_n}{a_{n-1}}$ where $0 < \rho \le 1$ and $P'(x_0)a_{n-1}(-1)^{n-1} \ge 0$.*
(6) *We have (with $a_0 = 1$)*

$$a_n a_{n-2} \le \sum_{j=0}^{n-1} |a_j| \left(n \left|\frac{a_{n-1}}{a_{n-2}}\right|\right)^{n-j} |a_{n-2}|.$$

(7) *We have (with $a_0 = 1$)*

$$|a_n| \le \sum_{j=0}^{n-1} |a_j| \left(2n^2\left(\min\left(\left|\frac{a_{n-1}}{a_{n-2}}\right|, \left|\frac{a_{n-1}}{a_{n-3}}\right|^{\frac{1}{2}}\right) + \min\left(\left|\frac{a_{n-2}}{a_{n-3}}\right|, \left|\frac{a_{n-2}}{a_{n-4}}\right|^{\frac{1}{2}}\right)\right)\right)^{n-j}.$$

PROOF. (1) and (2) are immediate corollaries of Rolle's theorem which states that the derivative f' of a real valued function f, which is continuous on a compact interval $[a, b]$ and differentiable on (a, b) with $f(a) = f(b)$, vanishes at at least one point in (a, b).

To (3): Suppose $P'(y_0) = 0$ and $P(y_0) \neq 0$. Then, by (1), y_0 is lying strictly between two roots $x_1 < x_2$ of P, and no other root of P' lies between x_1 and x_2. Therefore, either $P(y_0) > 0$ and $P''(y_0) < 0$ (local maximum), or $P(y_0) < 0$ and $P''(y_0) > 0$ (local minimum).

To (4): Assume that (4) is false. We choose $3 \leq i \leq n$ such that $a_{i-2} = a_{i-1} = 0$ and $a_i \neq 0$. Consider the hyperbolic polynomial

$$Q(x) = P^{(n-i)}(x) = b_0 x^i - b_1 x^{i-1} + \cdots + b_{i-3}x^3 + (-1)^i (n-i)! a_i$$

with $b_j \in \mathbb{R}$. Then $Q'(0) = 0$ and $Q(0) \neq 0$, but $Q(0)Q''(0) = 0$, contradicting (3).

To (5): We use Vieta's formulas to show the existence of a root of the form $n\rho \frac{a_n}{a_{n-1}}$ with $0 < \rho \leq 1$. If one root equals 0, then $a_n = 0$, and the existence is trivial. Suppose that no root vanishes. We can assume without loss of generality that $n\frac{a_n}{a_{n-1}} > 0$ (otherwise replace x by $-x$). For contradiction suppose there is no root of P in the interval $\left[0, n\frac{a_n}{a_{n-1}}\right]$. It is not possible that all roots are negative, since a_n and a_{n-1} have the same sign. So there are roots $x_{j_1}, \ldots, x_{j_k} > n\frac{a_n}{a_{n-1}}$. For a fixed $1 \leq i \leq k$ we have

$$x_{j_i} > n\frac{a_n}{a_{n-1}} = n\frac{x_1 \cdots x_n}{x_2 \cdots x_n + \cdots + x_1 \cdots x_{n-1}},$$

and thus

$$\frac{x_{j_i}}{x_1} + \cdots + \frac{x_{j_i}}{x_n} > n.$$

This inequality is only weakened, if one leaves away the negative terms:

$$\frac{x_{j_i}}{x_{j_1}} + \cdots + \frac{x_{j_i}}{x_{j_k}} > n.$$

But then there must exist some $l \in \{1, \ldots, k\}\backslash\{i\}$ such that $x_{j_i} > x_{j_l}$. And since i was arbitrary, it leads to a contradiction. Therefore the existence follows.

From all such roots choose one x_0 of minimal absolute-value. Then, either $x_0 = 0$, or, if $x_0 \neq 0$, then P has the same sign inside the segment with endpoints 0 and x_0. In both cases we find that

$$0 \geq P'(x_0)x_0 P(0) = P'(x_0)x_0(-1)^n a_n = (-1)^n P'(x_0) n\rho \frac{a_n^2}{a_{n-1}}$$

which is equivalent to $P'(x_0)a_{n-1}(-1)^{n-1} \geq 0$. (If $x_0 = 0$ the desired inequality is evident.)

To (6): The inequality is clearly satisfied, if $a_{n-2} = 0$. So let us suppose that $a_{n-2} \neq 0$. Consider the hyperbolic polynomial

$$P'(x) = nx^{n-1} - (n-1)a_1 x^{n-2} + \cdots + (-1)^{n-2} 2a_{n-2}x + (-1)^{n-1}a_{n-1}.$$

Use (5) to see that P' has a root $y_0 = (n-1)\rho\frac{a_{n-1}}{2a_{n-2}}$, with $0 < \rho \leq 1$, such that

$$P''(y_0)a_{n-2}(-1)^{n-2} \geq 0.$$

If $P(y_0) = 0$, then

$$\begin{aligned} |a_n| &\le |y_0|^n + |a_1||y_0|^{n-1} + \cdots + |a_{n-1}||y_0| \\ &= \sum_{j=0}^{n-1} |a_j| \left(\frac{n-1}{2} \rho \left| \frac{a_{n-1}}{a_{n-2}} \right| \right)^{n-j} \\ &\le \sum_{j=0}^{n-1} |a_j| \left(n \left| \frac{a_{n-1}}{a_{n-2}} \right| \right)^{n-j}, \end{aligned}$$

where $a_0 = 1$, from which the statement follows.

If $P(y_0) \neq 0$, then (3) implies $P''(y_0)P(y_0) < 0$. Therefore, $P''(y_0) \neq 0$. In the case that $P''(y_0) > 0$, we have $(-1)^{n-2}a_{n-2} > 0$ and $P(y_0) < 0$. Thus, multiplying the inequality

$$0 > P(y_0) = y_0^n - a_1 y_0^{n-1} + \cdots + (-1)^n a_n$$

by $(-1)^{n-2}a_{n-2}$ gives

$$\begin{aligned} a_n a_{n-2} &< (-1)^{n-2} a_{n-2} \left(-y_0^n + a_1 y_0^{n-1} + \cdots + (-1)^n a_{n-1} \right) \\ &\le \sum_{j=0}^{n-1} |a_j| \left(n \left| \frac{a_{n-1}}{a_{n-2}} \right| \right)^{n-j} |a_{n-2}|. \end{aligned}$$

In the case where $P''(y_0) < 0$, we have $(-1)^{n-2}a_{n-2} < 0$ and $P(y_0) > 0$. In an analogous way we obtain the desired inequality.

To (7): For a root x of P we obviously have

$$|a_n| \le \sum_{j=0}^{n-1} |a_j||x|^{n-j}.$$

Let us apply lemma 5.4.1 to the polynomial P'. Suppose the roots y_i of P' satisfy the inequalities $|y_1| \le |y_2| \le \cdots \le |y_{n-1}|$. Then

$$\begin{aligned} |y_2| &\le 2(n-1)^2 \left(\min\left(\left| \frac{a_{n-1}}{2a_{n-2}} \right|, \left| \frac{a_{n-1}}{3a_{n-3}} \right|^{\frac{1}{2}} \right) + \min\left(\left| \frac{2a_{n-2}}{3a_{n-3}} \right|, \left| \frac{2a_{n-2}}{4a_{n-4}} \right|^{\frac{1}{2}} \right) \right) \\ &\le 2n^2 \left(\min\left(\left| \frac{a_{n-1}}{a_{n-2}} \right|, \left| \frac{a_{n-1}}{a_{n-3}} \right|^{\frac{1}{2}} \right) + \min\left(\left| \frac{a_{n-2}}{a_{n-3}} \right|, \left| \frac{a_{n-2}}{a_{n-4}} \right|^{\frac{1}{2}} \right) \right). \end{aligned}$$

Since, by (2), between y_1 and y_2 there is a root x of P with $|x| \le |y_2|$, the required inequality follows. □

5.5. Local boundedness of the derivatives of the roots

With the preliminary work of the previous section we are now able to show the local boundedness of the derivatives of the roots of a C^n curve of hyperbolic polynomials of degree n. The essential part of this proof is the following lemma.

5.5.1. Lemma. *Let the polynomial*

$$P(t)(x) = \sum_{j=0}^{m-r-1} B_j(t) x^{m-j} + \sum_{j=0}^{r} A_j(t) x^{r-j},$$

be hyperbolic for all $t \in [-1,1]$. Suppose that all B_i are bounded functions on $[-1,1]$, and that all A_i are functions of class C^i on $[-1,1]$, respectively. Let $A_0(t) \neq 0$, for all $t \in [-1,1]$, $A_{r-1}(0) \neq 0$, and $A_r(0) = 0$. Then, for some constant $C > 0$, depending only on the degree of the polynomial P,

$$\left| \frac{A_r'(0)}{A_{r-1}(0)} \right| \le \left(\sup_{i,t} |B_i(t)| + \max_{i,t,j\le i} |A_i^{(j)}(t)| + \max_t |A_0(t)|^{-1} + 2 \right)^C. \tag{5.15}$$

REMARK. For the following consideration let us assume that all coefficients of P in the above lemma are of class C^m on $[-1,1]$ and that $B_0 = 1$. The conditions $A_{r-1}(0) \neq 0$ and $A_r(0) = 0$ mean that 0 is a simple root of $P(0)$. By the splitting lemma 4.2.3, we may factorize $P(t) = P_1(t) \cdot P_2(t)$ near $t = 0$, where $P_1(t)(x) = x - C_1(t)$ and $P_2(t)(x) = x^{m-1} - D_1(t)x^{m-2} + \dots + (-1)^{m-1}D_{m-1}(t)$ with functions $C_1, D_1, \dots, D_{m-1}$ of class C^m, $C_1(0) = 0$, and $D_{m-1}(0) \neq 0$. Consequently, we have $A_r(t) = (-1)^m C_1(t) D_{m-1}(t)$ and $A_{r-1}(t) = (-1)^{m-1} C_1(t) D_{m-2}(t) + (-1)^{m-1} D_{m-1}(t)$, whence

$$\left| \frac{A_r'(0)}{A_{r-1}(0)} \right| = \left| \frac{C_1'(0) D_{m-1}(0)}{D_{m-1}(0)} \right| = |C_1'(0)|.$$

So, under the above assumptions, the inequality (5.15) may be interpreted as an estimate of the derivative at $t = 0$ which belongs to the single root 0 of $P(0)$ in terms of the coefficients of $P(t)$ and its derivatives up to order r.

The proof of lemma 5.5.1 will run from 5.5.2 till 5.5.11.

5.5.2. We introduce the following notation: $a_i = A_i(0)$ and $a_i^{(j)} = A_i^{(j)}(0)$. Next, we choose $r+1$ positive numbers $M_0 < M_1 < \dots < M_r$, sufficiently large that all estimates to come in this proof are fulfilled. For example, it is possible to set

$$M_0 = \left(m^m \left(\sup_{i,t} |B_i(t)| + \max_{i,t,j \le i} |A_i^{(j)}(t)| + \max_t |A_0(t)|^{-1} + 4 \right) \right)^{2m}$$

and

$$M_i = M_{i-1}^{4(m+4)^2} \qquad (1 \le i \le r).$$

Let I be the set of indices $1 \le i \le r-1$ satisfying the following system of conditions:

(I.1) $a_i a_{i-1} \neq 0$ and $\left| \frac{a_i}{a_{i-1}} \right| \le M_i$

(I.2) $\frac{1}{2} \le \frac{A_i(t)}{a_i} \le 2$, if $|t| \le M_i^{-1} \left| \frac{a_i}{a_{i-1}} \right|$

(I.3) $M_0^{-1} \le \frac{A_{i-1}(t)}{a_{i-1}} \le M_0$, if $|t| \le M_i^{-1} \left| \frac{a_i}{a_{i-1}} \right|$

(I.4) $|A_j(t)| \left| \frac{a_i}{a_{i-1}} \right|^{i-j} \le M_i |a_i|$, if $|t| \le M_i^{-1} \left| \frac{a_i}{a_{i-1}} \right|$ and $0 \le j \le i+1$.

And let II be the set of indices $i \in \{2, \dots, r-1\}$ satisfying:

(II.1) $a_i a_{i-2} \neq 0$ and $\left| \frac{a_i}{a_{i-2}} \right| \le M_i$

(II.2) $\frac{1}{2} \le \frac{A_i(t)}{a_i} \le 2$, if $|t| \le M_i^{-1} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{1}{2}}$

(II.3) $M_0^{-1} \le \frac{A_{i-2}(t)}{a_{i-2}} \le M_0$, if $|t| \le M_i^{-1} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{1}{2}}$

(II.4) $|A_j(t)| \left| \frac{a_i}{a_{i-2}} \right|^{\frac{i-j}{2}} \le M_i |a_i|$, if $|t| \le M_i^{-1} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{1}{2}}$ and $0 \le j \le i-1$.

To shorten notation let us write $(\mathrm{I.4})_{j_0}$ and $(\mathrm{II.4})_{j_0}$ for the conditions (I.4) and (II.4) with $j = j_0$, respectively. Define $J = \mathrm{I} \cup \mathrm{II}$. Note that r cannot be in J, since $a_r = 0$, by assumption, contradicting (I.1) and (II.1).

5.5.3. *J* **is not empty.** More precisely: *We claim that* $1 \in \mathrm{I}$, *if* $|a_0 a_2| \le a_1^2$, *and* $2 \in \mathrm{II}$, *if* $|a_0 a_2| \ge a_1^2$.

For the hyperbolic polynomial

$$\left(\frac{\partial}{\partial x} \right)^{r-2} P(t)(x) = \sum_{j=0}^{m-r-1} \frac{(m-j)!}{(m-r-j+2)!} B_j(t) x^{m-r-j+2} + \frac{r!}{2!} A_0(t) x^2 + \frac{(r-1)!}{1!} A_1(t) x + (r-2)! A_2(t),$$

we use lemma 5.4.4(6) and obtain

$$\begin{aligned}
A_2(t)A_0(t) &\le m^{m^2}\left(|B_0(t)|\left|\frac{A_1(t)}{A_0(t)}\right|^{m-r+2} + \cdots + |B_{m-r-1}(t)|\left|\frac{A_1(t)}{A_0(t)}\right|^3\right.\\
&\qquad\left. + |A_0(t)|\left|\frac{A_1(t)}{A_0(t)}\right|^2 + |A_1(t)|\left|\frac{A_1(t)}{A_0(t)}\right|\right)|A_0(t)|\\
&\le M_0^{\frac{1}{2}}A_1^2(t).
\end{aligned}$$

On the other hand,

$$\frac{a_0}{A_0(t)} \le \frac{|a_0|}{|A_0(t)|} \le \max_t |A_0(t)| \cdot \max_t |A_0(t)|^{-1} \le M_0^{\frac{1}{2}},$$

whence, for all $t \in [-1,1]$,

$$A_2(t)a_0 \le M_0^{\frac{1}{2}} A_2(t)A_0(t) \le M_0 A_1^2(t). \tag{5.16}$$

Now, consider the case that $|a_0a_2| \le a_1^2$. Then, $a_1 \neq 0$, for otherwise $a_2 = a_1 = 0$, since $a_0 \neq 0$ by assumption, which contradicts lemma 5.4.4(4). We put $t = \pm M_0^{-1}a_1$ into (5.16) and use Taylor's formula on both sides of the inequality:

$$a_2a_0 \pm a_2^{(1)}M_0^{-1}a_1a_0 + \frac{A_2^{(2)}(\xi)}{2!}M_0^{-2}a_1^2a_0 \le M_0\left(a_1 \pm A_1^{(1)}(\eta)M_0^{-1}a_1\right)^2.$$

This implies

$$\begin{aligned}
\pm a_2^{(1)}M_0^{-1}a_1a_0 &\le M_0a_1^2 \pm 2A_1^{(1)}(\eta)a_1^2 + (A_1^{(1)}(\eta))^2M_0^{-1}a_1^2\\
&\qquad - a_2a_0 - \frac{A_2^{(2)}(\xi)}{2!}M_0^{-2}a_1^2a_0\\
&\le M_0a_1^2 + 2|A_1^{(1)}(\eta)|a_1^2 + (A_1^{(1)}(\eta))^2M_0^{-1}a_1^2\\
&\qquad + |a_2a_0| + \frac{|A_2^{(2)}(\xi)|}{2!}M_0^{-2}a_1^2|a_0|\\
&\le (M_0 + 2M_0 + M_0 + 1 + 1)a_1^2 \qquad (\text{since } |a_0a_2| \le a_1^2)\\
&\le M_0^2a_1^2,
\end{aligned}$$

whence

$$|a_2^{(1)}| \le M_0^3|a_0|^{-1}|a_1| \le M_0^4|a_1|.$$

Thus, for $|t| \le M_1^{-1}\left|\frac{a_1}{a_0}\right|$, we find

$$\begin{aligned}
|A_2(t)a_0| &\le |a_2a_0| + |a_2^{(1)}||t||a_0| + \frac{|A_2^{(2)}(\xi)|}{2!}|t|^2|a_0|\\
&\le a_1^2 + M_0^4|a_1|M_1^{-1}\left|\frac{a_1}{a_0}\right||a_0| + M_0M_1^{-2}\left|\frac{a_1}{a_0}\right|^2|a_0|\\
&\le M_1a_1^2.
\end{aligned}$$

The conditions of I are satisfied for $i = 1$, and so $1 \in J$: Condition (I.1) is clear, since $\left|\frac{a_1}{a_0}\right| \le M_0^2 \le M_1$. To see (I.2) observe that, for $|t| \le M_1^{-1}\left|\frac{a_1}{a_0}\right|$,

$$\frac{1}{2} \le 1 - \frac{|A_1^{(1)}(\xi)|}{|a_1|}|t| \le \frac{A_1(t)}{a_1} = 1 + \frac{A_1^{(1)}(\xi)}{a_1}t \le 1 + \frac{|A_1^{(1)}(\xi)|}{|a_1|}|t| \le 2,$$

since $\frac{|A_1^{(1)}(\xi)|}{|a_1|}|t| \le M_0M_1^{-1}|a_0|^{-1} \le M_0^2M_1^{-1} \le \frac{1}{2}$. Condition (I.3) follows from the definition of M_0, and it implies $(\mathrm{I.4})_0$. Condition (I.2) implies $(\mathrm{I.4})_1$, and $(\mathrm{I.4})_2$ was shown above.

In the second case, when $|a_0a_2| \ge a_1^2$, we find as before that $a_2 \neq 0$. We put $t = \pm M_0^{-1}\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}}$ into (5.16), and we use Taylor's formula on both sides of the inequality:

$$a_2a_0 \pm a_2^{(1)}M_0^{-1}\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}}a_0 + \frac{A_2^{(2)}(\xi)}{2!}M_0^{-2}\left|\frac{a_2}{a_0}\right|a_0 \le M_0\left(a_1 \pm A_1^{(1)}(\eta)M_0^{-1}\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}}\right)^2.$$

This implies

$$\begin{aligned}
\pm a_2^{(1)}M_0^{-1}\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}}a_0 &\le M_0a_1^2 \pm 2A_1^{(1)}(\eta)a_1\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}} + (A_1^{(1)}(\eta))^2M_0^{-1}\left|\frac{a_2}{a_0}\right| \\
&\quad - a_2a_0 - \frac{A_2^{(2)}(\xi)}{2!}M_0^{-2}\left|\frac{a_2}{a_0}\right|a_0 \\
&\le M_0a_1^2 + 2|A_1^{(1)}(\eta)||a_1|\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}} + (A_1^{(1)}(\eta))^2M_0^{-1}\left|\frac{a_2}{a_0}\right| \\
&\quad + |a_2a_0| + \frac{|A_2^{(2)}(\xi)|}{2!}M_0^{-2}\left|\frac{a_2}{a_0}\right||a_0| \\
&\le (M_0^2 + 2M_0 + M_0^2 + M_0 + M_0)|a_2| \quad (\text{since } a_1^2 \le |a_0a_2|) \\
&\le M_0^3|a_2|,
\end{aligned}$$

whence

$$|a_2^{(1)}| \le M_0^4|a_0|^{-\frac{1}{2}}|a_2|^{\frac{1}{2}} \le M_0^5|a_2|^{\frac{1}{2}}.$$

Consequently,

$$\frac{1}{2} \le \frac{A_2(t)}{a_2} \le 2,$$

for $|t| \le M_2^{-1}\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}}$, since $\frac{A_2(t)}{a_2} = 1 + \frac{a_2^{(1)}}{a_2}t + \frac{A_2^{(2)}(\xi)}{2a_2}t^2$ and

$$\frac{|a_2^{(1)}|}{|a_2|}|t| + \frac{|A_2^{(2)}(\xi)|}{2|a_2|}|t|^2 \le M_0^5M_2^{-1}|a_0|^{-\frac{1}{2}} + M_0M_2^{-2}|a_0|^{-1} \le \frac{1}{2}.$$

Further, if $|t| \le M_2^{-1}\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}}$, then

$$\begin{aligned}
|A_1(t)| &\le |a_1| + |A_1^{(1)}(\xi)||t| \le |a_0a_2|^{\frac{1}{2}} + M_0M_2^{-1}\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}} \\
&\le |a_0a_2|^{\frac{1}{2}}\left(1 + M_0^2M_2^{-1}\right) \le 2|a_0a_2|^{\frac{1}{2}}.
\end{aligned}$$

The index $i = 2$ satisfies the conditions of II, and so $2 \in J$: Condition (II.1) is clear, and (II.2) was shown above. Condition (II.3) follows from the definition of M_0, and it implies $(\text{II}.4)_0$. Finally, $(\text{II}.4)_1$ has been shown in the last computation. Therefore the proof of the claim is complete.

5.5.4. *We claim that $r - 1 \in J$.*

Suppose otherwise. Let $i < r - 1$ be the largest index belonging to J. Then $i + 2 \le r$. We assert the following implications:

(I′) If $i \in \mathrm{I}$ and $|a_ia_{i+2}| \le a_{i+1}^2$, then $i + 1 \in \mathrm{I}$.
(I″) If $i \in \mathrm{I}$ and $|a_ia_{i+2}| \ge a_{i+1}^2$, then $i + 1 \in \mathrm{II}$.
(II′) If $i \in \mathrm{II}$ and $|a_ia_{i+2}| \le a_{i+1}^2$, then $i + 1 \in \mathrm{I}$.
(II″) If $i \in \mathrm{II}$ and $|a_ia_{i+2}| \ge a_{i+1}^2$, then $i + 1 \in \mathrm{II}$.

5.5.5. Case I. Let us assume that i satisfies the conditions of I, without specifying the subcases (I′) and (I″). For $|t| \le M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right|$, we get, from $(\mathrm{I}.4)_{i+1}$ and $(\mathrm{I}.4)_0$,

$$\begin{aligned}\left|\sum_{j=0}^{i} \frac{a_{i+1}^{(j)}}{j!} t^j\right| &= \left|A_{i+1}(t) - \frac{A_{i+1}^{(i+1)}(\xi)}{(i+1)!} t^{i+1}\right| \le |A_{i+1}(t)| + \frac{|A_{i+1}^{(i+1)}(\xi)|}{(i+1)!} |t|^{i+1} \\ &\le M_i \left|\frac{a_i^2}{a_{i-1}}\right| + M_0 M_i^{-i-1} \left|\frac{a_i}{a_{i-1}}\right|^{i+1} \\ &\le M_i \left|\frac{a_i^2}{a_{i-1}}\right| + M_0^2 M_i^{-i} |a_i| \left|\frac{a_i}{a_{i-1}}\right| \\ &\le 2M_i \left|\frac{a_i^2}{a_{i-1}}\right|.\end{aligned}$$

We use lemma 5.4.2 and find

$$|a_{i+1}^{(j)}| \le M_0 M_i^{j+1} \left|\frac{a_i^{2-j}}{a_{i-1}^{1-j}}\right| \qquad (0 \le j \le i). \tag{5.17}$$

Consequently, for $|t| \le M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right|$, we obtain, with (5.17) and $(\mathrm{I}.4)_0$,

$$\begin{aligned}|A_{i+1}'(t)| &\le \sum_{j=0}^{i-1} \frac{|a_{i+1}^{(j+1)}|}{j!} |t|^j + \frac{|A_{i+1}^{(i+1)}(\xi)|}{i!} |t|^i \\ &\le \sum_{j=0}^{i-1} \frac{1}{j!} M_0 M_i^{j+2} \left|\frac{a_i^{2-(j+1)}}{a_{i-1}^{1-(j+1)}}\right| M_i^{-j} \left|\frac{a_i}{a_{i-1}}\right|^j + M_0 M_i^{-i} \left|\frac{a_i}{a_{i-1}}\right|^i \\ &\le \sum_{j=0}^{i-1} \frac{1}{j!} M_0 M_i^2 |a_i| + M_0^2 M_i^{-i+1} |a_i| \\ &\le M_0^2 M_i^2 |a_i|.\end{aligned} \tag{5.18}$$

For $|t| \le \frac{1}{2} M_0^{-2} M_i^{-2} \left|\frac{a_{i+1}}{a_i}\right| \left(\le \frac{1}{2} M_0^{-2} M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right| \le M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right|, \text{ by } (\mathrm{I}.4)_{i+1}\right)$, the estimate (5.18) implies

$$|A_{i+1}(t) - a_{i+1}| = |A_{i+1}'(\xi)||t| \le M_0^2 M_i^2 |a_i| \cdot \frac{1}{2} M_0^{-2} M_i^{-2} \left|\frac{a_{i+1}}{a_i}\right| = \frac{1}{2} |a_{i+1}|. \tag{5.19}$$

Consider the hyperbolic polynomial

$$\begin{aligned}\left(\frac{\partial}{\partial x}\right)^{r-(i+2)} P(t)(x) = &\sum_{j=0}^{m-r-1} \frac{(m-j)!}{(m-r+(i+2)-j)!} B_j(t) x^{m-r+(i+2)-j} \\ &+ \frac{r!}{(i+2)!} A_0(t) x^{i+2} + \cdots + (r-(i+2))! A_{i+2}(t).\end{aligned}$$

Applying lemma 5.4.4(7) we obtain

$$\begin{aligned}|A_{i+2}(t)| \le M_0 &\sum_{j=i+3}^{m} \left(\left|\frac{A_i(t)}{A_{i-1}(t)}\right| + \left|\frac{A_{i+1}(t)}{A_i(t)}\right|\right)^j \\ &+ M_0 \sum_{j=0}^{i+1} |A_j(t)| \left(\left|\frac{A_i(t)}{A_{i-1}(t)}\right| + \left|\frac{A_{i+1}(t)}{A_i(t)}\right|\right)^{i+2-j}.\end{aligned}$$

The inequalities (I.2), (I.3), and $(I.4)_{i+1}$ provide, for $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|$, the following estimates:

$$\left|\frac{A_i(t)}{A_{i-1}(t)}\right| \le 2M_0\left|\frac{a_i}{a_{i-1}}\right|$$

and

$$\left|\frac{A_{i+1}(t)}{A_i(t)}\right| \le M_i\left|\frac{a_i}{A_i(t)}\right|\left|\frac{a_i}{a_{i-1}}\right| \le 2M_i\left|\frac{a_i}{a_{i-1}}\right|. \tag{5.20}$$

Therefore we get, for $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|$,

$$|A_{i+2}(t)| \le M_0\sum_{j=i+3}^{m}\left(4M_i\left|\frac{a_i}{a_{i-1}}\right|\right)^j + M_0\sum_{j=0}^{i+1}|A_j(t)|\left(4M_i\left|\frac{a_i}{a_{i-1}}\right|\right)^{i+2-j}.$$

We estimate the first sum on the right-hand side, using $(I.4)_0$ and (I.1),

$$\begin{aligned}
\sum_{j=i+3}^{m}\left(4M_i\left|\frac{a_i}{a_{i-1}}\right|\right)^j &= \sum_{j=0}^{m-i-3}(4M_i)^{i+3+j}\left|\frac{a_i}{a_{i-1}}\right|^i\left|\frac{a_i}{a_{i-1}}\right|^{j+3}\\
&\le \sum_{j=0}^{m-i-3}(4M_i)^{i+3+j}M_0M_i|a_i|\left|\frac{a_i}{a_{i-1}}\right|^{j+3}\\
&\le M_0\sum_{j=0}^{m-i-3}(4M_i)^{i+3+j}M_i^{j+2}\left|\frac{a_i^3}{a_{i-1}^2}\right|\\
&\le M_0M_i^{2m}\left|\frac{a_i^3}{a_{i-1}^2}\right|.
\end{aligned}$$

Thus, for $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|$, we have, with $(I.4)_j$,

$$\begin{aligned}
|A_{i+2}(t)| &\le M_0^2M_i^{2m}\left|\frac{a_i^3}{a_{i-1}^2}\right| + M_0\sum_{j=0}^{i+1}|A_j(t)|\left(4M_i\left|\frac{a_i}{a_{i-1}}\right|\right)^{i+2-j}\\
&\le M_0^2M_i^{2m}\left|\frac{a_i^3}{a_{i-1}^2}\right| + M_0\sum_{j=0}^{i+1}M_i|a_i|\left|\frac{a_i}{a_{i-1}}\right|^{j-i}\left(4M_i\left|\frac{a_i}{a_{i-1}}\right|\right)^{i+2-j}\\
&\le M_i^{2m+2}\left|\frac{a_i^3}{a_{i-1}^2}\right|.
\end{aligned}$$

Using this estimate and $(I.4)_0$, we conclude in the same way as above that, for $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|$,

$$\begin{aligned}
\left|\sum_{j=0}^{i+1}\frac{a_{i+2}^{(j)}}{j!}t^j\right| &= \left|A_{i+2}(t) - \frac{A_{i+2}^{(i+2)}(\xi)}{(i+2)!}t^{i+2}\right| \le |A_{i+2}(t)| + \frac{|A_{i+2}^{(i+2)}(\xi)|}{(i+2)!}|t|^{i+2}\\
&\le M_i^{2m+2}\left|\frac{a_i^3}{a_{i-1}^2}\right| + M_0M_i^{-i-2}\left|\frac{a_i}{a_{i-1}}\right|^{i+2}\\
&\le M_i^{2m+2}\left|\frac{a_i^3}{a_{i-1}^2}\right| + M_0^2M_i^{-i-1}|a_i|\left|\frac{a_i}{a_{i-1}}\right|^2\\
&\le 2M_i^{2m+2}\left|\frac{a_i^3}{a_{i-1}^2}\right|.
\end{aligned}$$

Use again lemma 5.4.2 to obtain:

$$|a_{i+2}^{(j)}| \le M_0 M_i^{2m+2+j} \left| \frac{a_i^{3-j}}{a_{i-1}^{2-j}} \right| \qquad (0 \le j \le i+1). \tag{5.21}$$

Using (5.21) and $(\mathrm{I}.4)_0$ we find, for $|t| \le M_i^{-1} \left| \frac{a_i}{a_{i-1}} \right|$,

$$\begin{aligned}
|A''_{i+2}(t)| &\le \sum_{j=0}^{i-1} \left| a_{i+2}^{(j+2)} \right| |t|^j + \frac{|A_{i+2}^{(i+2)}(\xi)|}{i!} |t|^i \\
&\le \sum_{j=0}^{i-1} M_0 M_i^{2m+2+(j+2)} \left| \frac{a_i^{3-(j+2)}}{a_{i-1}^{2-(j+2)}} \right| M_i^{-j} \left| \frac{a_i}{a_{i-1}} \right|^j + M_0 M_i^{-i} \left| \frac{a_i}{a_{i-1}} \right|^i \\
&\le \sum_{j=0}^{i-1} M_0 M_i^{2m+4} |a_i| + M_0^2 M_i^{-i+1} |a_i| \\
&\le M_i^{2m+6} |a_i|.
\end{aligned} \tag{5.22}$$

Let us apply lemma 5.4.4(6) to the polynomial $\left(\frac{\partial}{\partial x}\right)^{r-(i+2)} P(t)(x)$:

$$A_{i+2}(t) A_i(t) \le \frac{1}{4} M_0 \left(\sum_{j=i+3}^{m} \left| \frac{A_{i+1}(t)}{A_i(t)} \right|^j + \sum_{j=0}^{i+1} |A_j(t)| \left| \frac{A_{i+1}(t)}{A_i(t)} \right|^{i+2-j} \right) |A_i(t)|.$$

By (I.2) and (5.20), we find that, for $|t| \le M_i^{-1} \left| \frac{a_i}{a_{i-1}} \right|$,

$$\begin{aligned}
A_{i+2}(t) a_i &\le M_0 \left(\sum_{j=i+3}^{m} \left| \frac{A_{i+1}(t)}{A_i(t)} \right|^j + \sum_{j=0}^{i+1} |A_j(t)| \left| \frac{A_{i+1}(t)}{A_i(t)} \right|^{i+2-j} \right) |a_i| \\
&\le M_0 \left| \frac{A_{i+1}(t)}{A_i(t)} \right|^2 |a_i| \\
&\quad \cdot \left(\sum_{j=i+1}^{m-2} \left(2M_i \left| \frac{a_i}{a_{i-1}} \right| \right)^j + \sum_{j=0}^{i} |A_j(t)| \left(2M_i \left| \frac{a_i}{a_{i-1}} \right| \right)^{i-j} \right) \\
&\quad + M_0 |A_{i+1}(t)| \left| \frac{A_{i+1}(t)}{A_i(t)} \right| |a_i|.
\end{aligned}$$

We estimate the first sum on the right-hand side, using $(\mathrm{I}.4)_0$ and (I.1):

$$\begin{aligned}
\sum_{j=i+1}^{m-2} (2M_i)^j \left| \frac{a_i}{a_{i-1}} \right|^j &= \sum_{j=0}^{m-i-3} (2M_i)^{i+1+j} \left| \frac{a_i}{a_{i-1}} \right|^i \left| \frac{a_i}{a_{i-1}} \right|^{j+1} \\
&\le \sum_{j=0}^{m-i-3} (2M_i)^{i+1+j} M_0 M_i |a_i| \left| \frac{a_i}{a_{i-1}} \right|^{j+1} \\
&\le \sum_{j=0}^{m-i-3} (2M_i)^{i+1+j} M_0 M_i^{j+2} |a_i| \\
&\le M_i^{2m} |a_i|.
\end{aligned}$$

Consequently, using $(\mathrm{I}.4)_j$ and (I.2), we obtain, for $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|$,

$$\begin{aligned}
A_{i+2}(t)a_i &\le M_0 \left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2 |a_i| \\
&\qquad \cdot \left(M_i^{2m}|a_i| + \sum_{j=0}^{i} M_i|a_i| \left|\frac{a_i}{a_{i-1}}\right|^{j-i} (2M_i)^{i-j} \left|\frac{a_i}{a_{i-1}}\right|^{i-j} \right) \\
&\qquad + M_0 \left|\frac{a_i}{A_i(t)}\right| |A_{i+1}(t)|^2 \\
&\le 2M_0 M_i^{2m} |A_{i+1}(t)|^2 \left|\frac{a_i}{A_i(t)}\right|^2 + M_0 \left|\frac{a_i}{A_i(t)}\right| |A_{i+1}(t)|^2 \\
&\le M_i^{2m+1}|A_{i+1}(t)|^2. \qquad (5.23)
\end{aligned}$$

All we have done till now is true in the case that $i \in \mathrm{I}$. In the following we want to consider separately the subcases (I′) and (I″).

5.5.6. Subcase (I′)**.** In the subcase (I′) we have $|a_i a_{i+2}| \le a_{i+1}^2$. Then $a_{i+1} \ne 0$ (otherwise $a_{i+1} = a_{i+2} = 0$, since (I.1) implies $a_i \ne 0$, a contradiction to lemma 5.4.4(4)).

By inequality (5.19) we have, for $|t| \le \frac{1}{2} M_0^{-2} M_i^{-2} \left|\frac{a_{i+1}}{a_i}\right|$,

$$\frac{1}{2} \le \frac{A_{i+1}(t)}{a_{i+1}} \le 2.$$

Therefore, condition (I.2) is satisfied for the index $i+1$. Setting $t = \pm M_i^{-3}\left|\frac{a_{i+1}}{a_i}\right|$ $\left(\le M_i^{-2}\left|\frac{a_i}{a_{i-1}}\right| \le M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|, \text{ by } (\mathrm{I}.4)_{i+1}\right)$ into (5.23), and using the previous result, we conclude with Taylor's formula:

$$a_i a_{i+2} \pm a_i a_{i+2}^{(1)} M_i^{-3} \left|\frac{a_{i+1}}{a_i}\right| + a_i \frac{A_{i+2}^{(2)}(\xi)}{2!} M_i^{-6} \left|\frac{a_{i+1}}{a_i}\right|^2 \le 4M_i^{2m+1}|a_{i+1}|^2.$$

Since $|a_i a_{i+2}| \le a_{i+1}^2$, this implies, with (5.22),

$$\begin{aligned}
\pm a_i a_{i+2}^{(1)} M_i^{-3} \left|\frac{a_{i+1}}{a_i}\right| &\le 4M_i^{2m+1}|a_{i+1}|^2 + |a_i a_{i+2}| + |a_i| \frac{|A_{i+2}^{(2)}(\xi)|}{2!} M_i^{-6} \left|\frac{a_{i+1}}{a_i}\right|^2 \\
&\le 4M_i^{2m+1}|a_{i+1}|^2 + |a_{i+1}|^2 + M_i^{2m}|a_{i+1}|^2 \\
&\le M_i^{2m+2}|a_{i+1}|^2,
\end{aligned}$$

whence

$$|a_{i+2}^{(1)}| \le M_i^{2m+5}|a_{i+1}|.$$

Thus, we have, for $|t| \le M_{i+1}^{-1}\left|\frac{a_{i+1}}{a_i}\right|$ $\left(\le M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right| \text{ as seen before}\right)$, using also (5.22),

$$\begin{aligned}
|A_{i+2}(t)a_i| &\le |a_i a_{i+2}| + |a_i||a_{i+2}^{(1)}||t| + |a_i| \frac{|A_{i+2}^{(2)}(\xi)|}{2!}|t|^2 \\
&\le |a_{i+1}|^2 + |a_i| M_i^{2m+5}|a_{i+1}| M_{i+1}^{-1}\left|\frac{a_{i+1}}{a_i}\right| \\
&\qquad + |a_i| M_i^{2m+6}|a_i| M_{i+1}^{-2}\left|\frac{a_{i+1}}{a_i}\right|^2 \\
&\le 2|a_{i+1}|^2. \qquad (5.24)
\end{aligned}$$

For $|t| \le M_{i+1}^{-1}\left|\frac{a_{i+1}}{a_i}\right|$ and $0 \le j \le i+1$, we get, from $(\mathrm{I}.4)_{i+1}$ and $(\mathrm{I}.4)_j$,

$$\begin{aligned} |A_j(t)|\left|\frac{a_{i+1}}{a_i}\right|^{i+1-j} &\le |A_j(t)|\left|\frac{a_{i+1}}{a_i}\right| M_i^{i-j}\left|\frac{a_i}{a_{i-1}}\right|^{i-j} \\ &\le M_i^{i+1-j}|a_{i+1}| \le M_{i+1}|a_{i+1}|. \end{aligned}$$

Now we are able to see that the index $i+1$ satisfies the conditions of I: Condition (I.1) is clear. We have already seen (I.2). Condition (I.3) holds true since $i \in \mathrm{I}$. Finally $(\mathrm{I}.4)_j$, for $0 \le j \le i+1$, has been shown in the previous computation, and $(\mathrm{I}.4)_{i+2}$ corresponds to (5.24).

5.5.7. Subcase (I$''$). Let us consider now the subcase (I$''$), where $|a_i a_{i+2}| \ge a_{i+1}^2$. Then, by lemma 5.4.4(4), $a_{i+2} \ne 0$.

Set $t = \pm M_i^{-2m-3}\left|\frac{a_{i+2}}{a_i}\right|^{\frac12}$ $\left(\le M_i^{-m-2}M_0^{\frac12}\left|\frac{a_i}{a_{i-1}}\right| \le M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|,\ \text{by } (5.21)_0\right)$ into (5.23) and use Taylor's formula on both sides of the inequality, then we get

$$\begin{aligned} a_i a_{i+2} \pm a_i a_{i+2}^{(1)} M_i^{-2m-3}\left|\frac{a_{i+2}}{a_i}\right|^{\frac12} + a_i \frac{A_{i+2}^{(2)}(\xi)}{2!} M_i^{-4m-6}\left|\frac{a_{i+2}}{a_i}\right| & \\ \le M_i^{2m+1}\left|a_{i+1} \pm A_{i+1}^{(1)}(\eta) M_i^{-2m-3}\left|\frac{a_{i+2}}{a_i}\right|^{\frac12}\right|^2&. \end{aligned}$$

This implies, with $a_{i+2}^2 \le |a_i a_{i+2}|$, (5.18), and (5.22),

$$\begin{aligned} &\pm a_i a_{i+2}^{(1)} M_i^{-2m-3}\left|\frac{a_{i+2}}{a_i}\right|^{\frac12} \\ &\le M_i^{2m+1}a_{i+1}^2 + 2|a_{i+1}||A_{i+1}^{(1)}(\eta)|M_i^{-2}\left|\frac{a_{i+2}}{a_i}\right|^{\frac12} + (A_{i+1}^{(1)}(\eta))^2 M_i^{-2m-5}\left|\frac{a_{i+2}}{a_i}\right| \\ &\quad + |a_i a_{i+2}| + |a_i|\frac{|A_{i+2}^{(2)}(\xi)|}{2!} M_i^{-4m-6}\left|\frac{a_{i+2}}{a_i}\right| \\ &\le \left(M_i^{2m+1} + 2M_0^2 + M_0^4 M_i^{-2m-1} + 1 + \frac12 M_i^{-2m}\right)|a_i a_{i+2}| \\ &\le 5M_i^{2m+1}|a_i a_{i+2}|. \end{aligned}$$

So we get

$$|a_{i+2}^{(1)}| \le 5M_i^{4m+4}|a_i a_{i+2}|^{\frac12}.$$

For $|t| \le M_i^{-4m-6}\left|\frac{a_{i+2}}{a_i}\right|^{\frac12}$, this gives, with (5.22):

$$\begin{aligned} |A_{i+2}(t) - a_{i+2}| = \left|a_{i+2}^{(1)}t + \frac{A_{i+2}^{(2)}(\xi)}{2!}t^2\right| &\le |a_{i+2}^{(1)}||t| + \frac{|A_{i+2}^{(2)}(\xi)|}{2!}|t|^2 \\ &\le 5M_i^{-2}|a_{i+2}| + \frac12 M_i^{-6m-6}|a_{i+2}| \le \frac12|a_{i+2}|, \end{aligned}$$

whence $i+2$ satisfies (II.2).

Finally, we obtain, from $(5.21)_0$ and $(\mathrm{I}.4)_j$,

$$\begin{aligned}
|A_j(t)| \left|\frac{a_{i+2}}{a_i}\right|^{\frac{i+2-j}{2}} &\le |A_j(t)| \left|\frac{a_{i+2}}{a_i}\right| \left(M_0 M_i^{2m+2} \left|\frac{a_i}{a_{i-1}}\right|^2\right)^{\frac{i-j}{2}} \\
&\le M_i |a_i| \left|\frac{a_i}{a_{i-1}}\right|^{j-i} \left|\frac{a_{i+2}}{a_i}\right| M_0^{\frac{i-j}{2}} M_i^{(m+1)(i-j)} \left|\frac{a_i}{a_{i-1}}\right|^{i-j} \\
&\le M_i^{(m+1)^2} |a_{i+2}|
\end{aligned}$$

which works for $|t| \le M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right|$ and $0 \le j \le i+1$.

Thus, the index $i+2$ satisfies the conditions of II: (II.1) and (II.2) are clear. Further, (II.3) is true, since $i \in \mathrm{I}$, and (II.4) has been shown in the last estimate.

5.5.8. Case II. Let us investigate now the case that i belongs to II. We use lemma 5.4.4(7) for the hyperbolic polynomial

$$\begin{aligned}
\left(\frac{\partial}{\partial x}\right)^{r-(i+1)} P(t)(x) = &\sum_{j=0}^{m-r-1} \frac{(m-j)!}{(m-r+(i+1)-j)!} B_j(t) x^{m-r+(i+1)-j} \\
&+ \frac{r!}{(i+1)!} A_0(t) x^{i+1} + \cdots + (r-(i+1))! A_{i+1}(t)
\end{aligned}$$

and obtain

$$\begin{aligned}
|A_{i+1}(t)| \le M_0 &\sum_{j=i+2}^{m} \left(\left|\frac{A_{i-1}(t)}{A_{i-2}(t)}\right| + \left|\frac{A_i(t)}{A_{i-2}(t)}\right|^{\frac{1}{2}}\right)^j \\
&+ M_0 \sum_{j=0}^{i} |A_j(t)| \left(\left|\frac{A_{i-1}(t)}{A_{i-2}(t)}\right| + \left|\frac{A_i(t)}{A_{i-2}(t)}\right|^{\frac{1}{2}}\right)^{i+1-j}.
\end{aligned}$$

For $|t| \le M_i^{-1} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$, we find, with $(\mathrm{II}.4)_{i-1}$ and (II.3),

$$\left|\frac{A_{i-1}(t)}{A_{i-2}(t)}\right| \le \frac{M_i |a_i a_{i-2}|^{\frac{1}{2}}}{|A_{i-2}(t)|} \le M_0 M_i \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \tag{5.25}$$

and, with (II.2) and (II.3),

$$\left|\frac{A_i(t)}{A_{i-2}(t)}\right|^{\frac{1}{2}} \le (2M_0)^{\frac{1}{2}} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}. \tag{5.26}$$

Using (5.25), (5.26), $(\mathrm{II}.4)_0$, and (II.1), we can estimate:

$$\begin{aligned}
\sum_{j=i+2}^{m} \left(\left|\frac{A_{i-1}(t)}{A_{i-2}(t)}\right| + \left|\frac{A_i(t)}{A_{i-2}(t)}\right|^{\frac{1}{2}}\right)^j &\le \sum_{j=i+2}^{m} (M_0^2 M_i)^j \left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}} \\
&\le \sum_{j=2}^{m-i} (M_0^2 M_i)^{i+j} M_0 M_i |a_i| \left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}} \\
&\le M_0 M_i |a_i| \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \sum_{j=2}^{m-i} (M_0^2 M_i)^{i+j} M_i^{\frac{j-1}{2}} \\
&\le M_0^{2m+2} M_i^{2m+1} |a_i| \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}.
\end{aligned}$$

Therefore, we get, for $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$, using (5.25), (5.26), (II.2) and (II.4)$_j$,

$$
\begin{aligned}
|A_{i+1}(t)| &\le M_0^{2m+3}M_i^{2m+1}|a_i|\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} + M_0\sum_{j=0}^{i}|A_j(t)|(M_0^2M_i)^{i+1-j}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{i+1-j}{2}} \\
&\le M_0^{2m+3}M_i^{2m+1}|a_i|\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} + M_0M_i|a_i|\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\sum_{j=0}^{i-1}(M_0^2M_i)^{i+1-j} \\
&\quad + 2M_0^3M_i|a_i|\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \\
&\le 3M_0^{2m+3}M_i^{2m+1}|a_i|\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \\
&\le M_i^{2m+2}|a_i|\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}.
\end{aligned}
\tag{5.27}
$$

Next we use lemma 5.4.4(7) for $\left(\frac{\partial}{\partial x}\right)^{r-(i+2)}P(t)(x)$, and we find the following estimate:

$$
\begin{aligned}
|A_{i+2}(t)| \le M_0\sum_{j=i+3}^{m}&\left(\left|\frac{A_{i+1}(t)}{A_i(t)}\right| + \left|\frac{A_i(t)}{A_{i-2}(t)}\right|^{\frac{1}{2}}\right)^j \\
&+ M_0\sum_{j=0}^{i+1}|A_j(t)|\left(\left|\frac{A_{i+1}(t)}{A_i(t)}\right| + \left|\frac{A_i(t)}{A_{i-2}(t)}\right|^{\frac{1}{2}}\right)^{i+2-j}.
\end{aligned}
$$

Considering (5.26) and the fact that, for $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$,

$$
\left|\frac{A_{i+1}(t)}{A_i(t)}\right| \le M_i^{2m+2}\left|\frac{a_i}{A_i(t)}\right|\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \le 2M_i^{2m+2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}, \tag{5.28}
$$

which follows from (5.27) and (II.2), we find

$$
\left|\frac{A_{i+1}(t)}{A_i(t)}\right| + \left|\frac{A_i(t)}{A_{i-2}(t)}\right|^{\frac{1}{2}} \le M_0M_i^{2m+2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}.
$$

So we may estimate, using also (II.4)$_0$ and (II.1):

$$
\begin{aligned}
\sum_{j=i+3}^{m}\left(\left|\frac{A_{i+1}(t)}{A_i(t)}\right| + \left|\frac{A_i(t)}{A_{i-2}(t)}\right|^{\frac{1}{2}}\right)^j &\le \sum_{j=i+3}^{m}(M_0M_i^{2m+2})^j\left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}} \\
&\le \sum_{j=3}^{m-i}(M_0M_i^{2m+2})^{i+j}M_0M_i|a_i|\left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}} \\
&\le M_0M_i|a_i|\left|\frac{a_i}{a_{i-2}}\right|\sum_{j=3}^{m-i}(M_0M_i^{2m+2})^{i+j}M_i^{\frac{j-2}{2}} \\
&\le M_0^{m+2}M_i^{2(m+1)^2}|a_i|\left|\frac{a_i}{a_{i-2}}\right|.
\end{aligned}
$$

The second sum on the right-hand side gives, with (II.2), $(\mathrm{II}.4)_j$, and (5.27),

$$\begin{aligned}
\sum_{j=0}^{i+1} |A_j(t)| \left(\left| \frac{A_{i+1}(t)}{A_i(t)} \right| + \left| \frac{A_i(t)}{A_{i-2}(t)} \right|^{\frac{1}{2}} \right)^{i+2-j} \\
&\le \sum_{j=0}^{i+1} |A_j(t)| (M_0 M_i^{2m+2})^{i+2-j} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{i+2-j}{2}} \\
&\le M_i |a_i| \left| \frac{a_i}{a_{i-2}} \right| \sum_{j=0}^{i-1} (M_0 M_i^{2m+2})^{i+2-j} + 2 M_0^2 M_i^{4m+4} |a_i| \left| \frac{a_i}{a_{i-2}} \right| \\
&\quad + M_0 M_i^{4m+4} |a_i| \left| \frac{a_i}{a_{i-2}} \right| \\
&\le M_i^{2(m+1)^2} |a_i| \left| \frac{a_i}{a_{i-2}} \right|.
\end{aligned}$$

So, for $|t| \le M_i^{-1} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{1}{2}}$, this implies:

$$|A_{i+2}(t)| \le M_i^{2(m+1)^2+1} |a_i| \left| \frac{a_i}{a_{i-2}} \right|. \tag{5.29}$$

From the inequalities (5.27) and (5.29) we can derive the following estimates, just applying lemma 5.4.2 as we did before:

$$|a_{i+1}^{(j)}| \le M_0 M_i^{2m+2+j} |a_i| \left| \frac{a_i}{a_{i-2}} \right|^{\frac{1-j}{2}} \qquad (0 \le j \le i) \tag{5.30}$$

and

$$|a_{i+2}^{(j)}| \le M_0 M_i^{2(m+1)^2+1+j} |a_i| \left| \frac{a_i}{a_{i-2}} \right|^{\frac{2-j}{2}} \qquad (0 \le j \le i+1). \tag{5.31}$$

For $|t| \le M_i^{-1} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{1}{2}}$, by (5.30) and $(\mathrm{II}.4)_0$, we obtain

$$\begin{aligned}
|A'_{i+1}(t)| &\le \sum_{j=0}^{i-1} |a_{i+1}^{(j+1)}| |t|^j + \frac{|A_{i+1}^{(i+1)}(\xi)|}{i!} |t|^i \\
&\le \sum_{j=0}^{i-1} M_0 M_i^{2m+2+j+1} |a_i| \left| \frac{a_i}{a_{i-2}} \right|^{-\frac{j}{2}} M_i^{-j} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{j}{2}} + M_0 M_i^{-i} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{i}{2}} \\
&\le \sum_{j=0}^{i-1} M_0 M_i^{2m+3} |a_i| + M_0^2 M_i^{-i+1} |a_i| \\
&\le \frac{1}{2} M_i^{2m+4} |a_i|.
\end{aligned} \tag{5.32}$$

For $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$, we get in a similar way, from (5.31) and $(\mathrm{II}.4)_0$,

$$\begin{aligned}
|A''_{i+2}(t)| &\le \sum_{j=0}^{i-1} |a_{i+2}^{(j+2)}||t|^j + \frac{|A_{i+2}^{(i+2)}(\xi)|}{i!}|t|^i \\
&\le \sum_{j=0}^{i-1} M_0 M_i^{2(m+1)^2+1+j+2}|a_i|\left|\frac{a_i}{a_{i-2}}\right|^{-\frac{j}{2}} M_i^{-j}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}} + M_0 M_i^{-i}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{i}{2}} \\
&\le \sum_{j=0}^{i-1} M_0 M_i^{2(m+1)^2+3}|a_i| + M_0^2 M_i^{-i+1}|a_i| \\
&\le M_i^{2(m+1)^2+4}|a_i|. \qquad (5.33)
\end{aligned}$$

Applying lemma 5.4.4(6) to $\left(\frac{\partial}{\partial x}\right)^{r-(i+2)} P(t)(x)$ and using (II.2), gives:

$$A_{i+2}(t)a_i \le M_0\left(\sum_{j=i+3}^{m}\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^j + \sum_{j=0}^{i+1}|A_j(t)|\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^{i+2-j}\right)|a_i|.$$

By (5.28), we have, for $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$,

$$\begin{aligned}
A_{i+2}(t)a_i \le M_0\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2 &\sum_{j=i+1}^{m-2}\left(2M_i^{2m+2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\right)^j |a_i| \\
&+ M_0\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2 \sum_{j=0}^{i}|A_j(t)|\left(2M_i^{2m+2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\right)^{i-j}|a_i| \\
&+ M_0|A_{i+1}(t)|\left|\frac{A_{i+1}(t)}{A_i(t)}\right||a_i|.
\end{aligned}$$

We may estimate, with $(\mathrm{II}.4)_0$ and (II.1):

$$\begin{aligned}
\sum_{j=i+1}^{m-2}\left(2M_i^{2m+2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\right)^j &= \sum_{j=1}^{m-i-2}(2M_i^{2m+2})^{i+j}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{i}{2}}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}} \\
&\le \sum_{j=1}^{m-i-2}(2M_i^{2m+2})^{i+j} M_0 M_i|a_i|\left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}} \\
&\le \sum_{j=1}^{m-i-2}(2M_i^{2m+2})^{i+j} M_0 M_i M_i^{\frac{j}{2}}|a_i| \\
&\le M_i^{2m^2+m}|a_i|.
\end{aligned}$$

Therefore, we obtain, for $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$, with $(\mathrm{II}.4)_j$ and (II.2):

$$\begin{aligned}
A_{i+2}(t)a_i &\le M_0 M_i^{2m^2+m}\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2 |a_i|^2 \\
&\quad + M_0\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2 |a_i| \sum_{j=0}^{i-1} M_i|a_i|\left|\frac{a_i}{a_{i-2}}\right|^{\frac{j-i}{2}}\left(2M_i^{2m+2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\right)^{i-j} \\
&\quad + M_0\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2 |a_i||A_i(t)| + M_0\left|\frac{a_i}{A_i(t)}\right||A_{i+1}(t)|^2 \\
&\le \left(4M_0M_i^{2m^2+m} + 4M_0^2M_i^{2m^2+2m+1} + 8M_0 + 2M_0\right)|A_{i+1}(t)|^2 \\
&\le M_i^{2(m+1)^2}|A_{i+1}(t)|^2. && (5.34)
\end{aligned}$$

As we did before, let us specify now the subcases (II′) and (II″).

5.5.9. Subcase (II′)**.** In the subcase (II′) we have $|a_i a_{i+2}| \le a_{i+1}^2$. Note that then $a_{i+1} \ne 0$ (otherwise lemma 5.4.4(4) is harmed).

For $|t| \le M_i^{-2m-4}\left|\frac{a_{i+1}}{a_i}\right| \left(\le M_0 M_i^{-2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \le M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}},\ \text{by } (5.30)_0\right)$, we find, by (5.32), that

$$|A_{i+1}(t) - a_{i+1}| = |A'_{i+1}(\xi)||t| \le \frac{1}{2}|a_{i+1}|, \tag{5.35}$$

which shows that $i+1$ satisfies condition (I.2).

Put $t = \pm M_i^{-2(m+1)^2}\left|\frac{a_{i+1}}{a_i}\right| \left(\le M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}},\ \text{by } (5.27)\right)$ into (5.34) and use Taylor's formula and (5.35):

$$\begin{aligned}
a_ia_{i+2} \pm a_ia_{i+2}^{(1)}M_i^{-2(m+1)^2}\left|\frac{a_{i+1}}{a_i}\right| + a_i\frac{A_{i+2}^{(2)}(\xi)}{2!}M_i^{-4(m+1)^2}\left|\frac{a_{i+1}}{a_i}\right|^2 \\
\le 4M_i^{2(m+1)^2}|a_{i+1}|^2.
\end{aligned}$$

This implies, since $|a_ia_{i+2}| \le a_{i+1}^2$ and by (5.33),

$$\begin{aligned}
\pm a_ia_{i+2}^{(1)}M_i^{-2(m+1)^2}\left|\frac{a_{i+1}}{a_i}\right| &\le 4M_i^{2(m+1)^2}|a_{i+1}|^2 + |a_{i+1}|^2 + M_i^{-2(m+1)^2+4}|a_{i+1}|^2 \\
&\le 6M_i^{2(m+1)^2}|a_{i+1}|^2,
\end{aligned}$$

whence

$$|a_{i+2}^{(1)}| \le 6M_i^{4(m+1)^2}|a_{i+1}|.$$

Thus, we have, for $|t| \le M_i^{-4(m+4)^2}\left|\frac{a_{i+1}}{a_i}\right| \left(\le M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|\right)$, using (5.33),

$$\begin{aligned}
|A_{i+2}(t)a_i| &\le |a_ia_{i+2}| + |a_i||a_{i+2}^{(1)}||t| + |a_i|\frac{|A_{i+2}^{(2)}(\xi)|}{2!}|t|^2 \\
&\le |a_{i+1}|^2 + |a_i|6M_i^{4(m+1)^2}|a_{i+1}|M_i^{-4(m+4)^2}\left|\frac{a_{i+1}}{a_i}\right| \\
&\quad + \frac{1}{2}|a_i|M_i^{2(m+1)^2+4}|a_i|M_i^{-8(m+4)^2}\left|\frac{a_{i+1}}{a_i}\right|^2 \\
&\le 2|a_{i+1}|^2. && (5.36)
\end{aligned}$$

For $|t| \le M_{i+1}^{-1} \left|\frac{a_{i+1}}{a_i}\right|$ and $0 \le j \le i-1$, we find, with (5.27) and $(\mathrm{II}.4)_j$,

$$\begin{aligned} |A_j(t)| \left|\frac{a_{i+1}}{a_i}\right|^{i+1-j} &\le |A_j(t)| \left|\frac{a_{i+1}}{a_i}\right| M_i^{(2m+2)(i-j)} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{i-j}{2}} \\ &\le M_i^{(2m+2)(i-j)+1} |a_{i+1}| \\ &\le M_{i+1}|a_{i+1}|. \end{aligned}$$

Now we are able to see that the index $i+1$ satisfies the conditions of I: Condition (I.1) is clear and (I.2) follows from (5.35). (I.3) is fulfilled, since $i \in \mathrm{II}$. Further, $(\mathrm{I}.4)_j$, for $0 \le j \le i-1$, has been shown just above, $(\mathrm{I}.4)_i$ and $(\mathrm{I}.4)_{i+1}$ are easy consequences of (I.3) and (I.2) (for $i+1$, respectively), and, finally, $(\mathrm{I}.4)_{i+2}$ corresponds to (5.36).

5.5.10. Subcase (II$''$). We consider now the subcase (II$''$), where $|a_i a_{i+2}| \ge a_{i+1}^2$. Then $a_{i+2} \neq 0$ (by lemma 5.4.4(4)).

Set $t = \pm M_i^{-2(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} \left(\le M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right|^{\frac{1}{2}}, \text{ by (5.29)}\right)$ into (5.34), and we find

$$\begin{aligned} a_i a_{i+2} \pm a_i a_{i+2}^{(1)} M_i^{-2(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} &+ a_i \frac{A_{i+2}^{(2)}(\xi)}{2!} M_i^{-4(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right| \\ &\le M_i^{2(m+1)^2} \left| a_{i+1} \pm A_{i+1}^{(1)}(\eta) M_i^{-2(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} \right|^2. \end{aligned}$$

This implies, by $a_{i+1}^2 \le |a_i a_{i+2}|$, (5.32), and (5.33),

$$\begin{aligned} &\pm a_i a_{i+2}^{(1)} M_i^{-2(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} \\ &\le M_i^{2(m+1)^2} a_{i+1}^2 + 2|a_{i+1}||A_{i+1}^{(1)}(\eta)| \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} \\ &\quad + (A_{i+1}^{(1)}(\eta))^2 M_i^{-2(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right| + |a_i a_{i+2}| + |a_i| \frac{|A_{i+2}^{(2)}(\xi)|}{2!} M_i^{-4(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right|, \\ &\le \left(M_i^{2(m+1)^2} + M_i^{2m+4} + \frac{1}{4} M_i^{-2m^2+6} + 1 + \frac{1}{2} M_i^{-2(m+1)^2+4}\right) |a_i a_{i+2}| \\ &\le 5 M_i^{2(m+1)^2} |a_i a_{i+2}|. \end{aligned}$$

So we get

$$|a_{i+2}^{(1)}| \le 5 M_i^{4(m+1)^2} |a_i a_{i+2}|^{\frac{1}{2}}.$$

For $|t| \le M_{i+2}^{-1} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}}$, this gives, using (5.33),

$$\begin{aligned} |A_{i+2}(t) - a_{i+2}| &\le |a_{i+2}^{(1)}||t| + \frac{|A_{i+2}^{(2)}(\xi)|}{2!} |t|^2 \\ &\le 5 M_i^{4(m+1)^2} M_{i+2}^{-1} |a_{i+2}| + \frac{1}{2} M_i^{2(m+1)^2+4} M_{i+2}^{-2} |a_{i+2}| \\ &\le \frac{1}{2} |a_{i+2}|, \end{aligned}$$

whence $i+2$ satisfies (II.2).

Finally, we investigate, whether $i+2$ fulfills (II.4):

$$\begin{aligned}
|A_j(t)|\left|\frac{a_{i+2}}{a_i}\right|^{\frac{i+2-j}{2}} &= |A_j(t)|\left|\frac{a_{i+2}}{a_i}\right|\left|\frac{a_{i+2}}{a_i}\right|^{\frac{i-j}{2}} \\
&\le |A_j(t)|\left|\frac{a_{i+2}}{a_i}\right|\left(M_i^{2(m+1)^2+1}\left|\frac{a_i}{a_{i-2}}\right|\right)^{\frac{i-j}{2}} \\
&\le M_i^{\frac{1}{2}(2(m+1)^2+1)(i-j)+1}|a_{i+2}| \\
&\le M_{i+2}|a_{i+2}|,
\end{aligned}$$

by (5.29) and $(\mathrm{II}.4)_j$, which works for $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$ and $0 \le j \le i-1$. For $j=i$, the statement follows from (II.2), and, for $j=i+1$, we observe that, if $|t| \le M_{i+2}^{-1}\left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}}$,

$$\begin{aligned}
|A_{i+1}(t)| &\le |a_{i+1}| + |A_{i+1}^{(1)}(\xi)||t| \\
&\le |a_i a_{i+2}|^{\frac{1}{2}} + \frac{1}{2}M_i^{2m+4}M_{i+2}^{-1}|a_i a_{i+2}|^{\frac{1}{2}} \\
&\le 2|a_i a_{i+2}|^{\frac{1}{2}},
\end{aligned}$$

since $a_{i+1}^2 \le |a_i a_{i+2}|$ and by (5.32). Thus, we have just shown that the index $i+2$ belongs to II, since (II.1) and (II.3) are obvious.

5.5.11. End of proof. We have supposed that i, being the largest index belonging to J, is strictly smaller than $r-1$. And we have seen that, consequently, either $i+1$ or $i+2$ belongs to J. Thus, $i \in \{r-2, r-1\}$. Suppose $r-2 \in J$. By the assumptions of the lemma, $0 = a_{r-2}a_r < a_{r-1}^2$. So the primed cases, (I′) or (II′), occur, whence $r-1 \in J$. This completes the proof of the claim in 5.5.4.

If $r-1$ satisfies the conditions of I, then inequality (5.17), with $j=1$, provides the estimate we are looking for. If $r-1$ satisfies the conditions of II, then inequality (5.30), with $j=1$, does it. Therefore, lemma 5.5.1 is proved.

5.5.12. LEMMA. *The preceding lemma 5.5.1 remains valid, if the inequality $A_{r-1}(0) \ne 0$ is replaced by the relations $A_{r-2}(0) \ne 0$ and $A_{r-1}(0) = 0$ and* (5.15) *is replaced by*

$$\left|\frac{A_r''(0)}{A_{r-2}(0)}\right| \le M_r.$$

PROOF. Rereading the proof of lemma 5.5.1 shows that now $r-2 \in J$. If $r-2$ belongs to I, then, by (5.22), we get

$$|A_r''(0)| \le M_{r-2}^{2m+6}|A_{r-2}(0)| \le M_r|A_{r-2}(0)|.$$

If $r-2$ belongs to I, then (5.33) gives

$$|A_r''(0)| \le M_{r-2}^{2(m+1)^2+4}|A_{r-2}(0)| \le M_r|A_{r-2}(0)|.$$

Therefore, the proof is complete. □

Finally, we can formulate and prove the main result of this chapter.

5.5.13. THEOREM. *Suppose that the polynomial*

$$P(t,y)(x) = x^n - a_1(t,y)x^{n-1} + \cdots + (-1)^n a_n(t,y)$$

is hyperbolic for any $(t,y) \in (-1,1)\times\mathcal{M}$, where $\mathcal{M}$ is a compact Hausdorff topological space, and the multiplicity of its roots does not exceed k. Furthermore, suppose that all partial derivatives $\frac{\partial^i}{\partial t^i}a_j(t,y)$ $(0 \le i \le k; 1 \le j \le n)$ are continuous functions on $(-1,1)\times\mathcal{M}$. Then, for any compact subset $K \subseteq (-1,1)\times\mathcal{M}$, there *exists a positive constant C_K such that, for all (differentiably chosen) roots $x_j(t,y)$ $(1 \le j \le n)$ of P, we have the following estimate*

$$\left|\frac{\partial}{\partial t}x_j(t,y)\right| < C_K \quad \textit{for all } (t,y) \in K.$$

PROOF. Note that, if $k=1$, all roots are simple all the time, and the statement follows easily from the implicit function theorem.

Suppose for contradiction that $\frac{\partial}{\partial t}x_j(t,y)$ is unbounded on a compact $K \subseteq (-1,1)\times\mathcal{M}$ for some $1 \le j \le n$. Without loss of generality let $j=1$, and we assume there is a sequence $(t_p,y_p)_{p\in\mathbb{N}} \subseteq K$ such that

$$\begin{aligned}
(t_p,y_p) &\overset{p\to\infty}{\longrightarrow} (t_\infty,y_\infty) \in K\\
x_1(t_p,y_p) &\overset{p\to\infty}{\longrightarrow} x_1(t_\infty,y_\infty)\\
\left|\frac{\partial}{\partial t}x_1(t_p,y_p)\right| &\overset{p\to\infty}{\longrightarrow} \infty.
\end{aligned}$$

Passing to a subsequence (again denoted by $(t_p,y_p)_p$) we may assume that the multiplicity of $x_1(t_p,y_p)$ equals, say, q for any $p\in\mathbb{N}$. Then, $1\le q\le k$. Therefore, the multiplicity s of $x_1(t_\infty,y_\infty)$ satisfies $q \le s \le k$; for, if $s<q$, then the sequence $(x_1(t_p,y_p))_p \subseteq \mathbb{R}$ had more than one limit, a contradiction.

Define

$$\begin{aligned}
Q_p(t)(\tilde{x}) &= P(t,y_p)(\tilde{x}+x_1(t_p,y_p))\\
&= \tilde{x}^n + b_{p,1}(t)\tilde{x}^{n-1}+\cdots+b_{p,n}(t),
\end{aligned}$$

where the coefficients $1=b_{p,0},b_{p,1},\dots,b_{p,n}$ take the following form, by Taylor's formula,

$$b_{p,j}(t) = \frac{1}{(n-j)!}\left(\frac{\partial}{\partial x}\right)^{n-j}\Bigg|_{x=x_1(t_p,y_p)} P(t,y_p)(x).$$

Remember that, by theorem 5.2.1, $\frac{\partial}{\partial t}x_1(t_p,y_p)$ has to satisfy, for all $p\in\mathbb{N}$, the following hyperbolic equation:

$$T_p(x) = b_{p,n-q}(t_p)x^q + \frac{1}{1!}b^{(1)}_{p,n-q+1}(t_p)x^{q-1}+\cdots+\frac{1}{q!}b^{(q)}_{p,n}(t_p) = 0.$$

Our goal is to show that all coefficients of $(b_{p,n-q}(t_p))^{-1}T_p(x)$ are bounded in p (that $b_{p,n-q}(t_p)\neq 0$ will be checked below). This would contradict the assumption that $\frac{\partial}{\partial t}x_1(t_p,y_p)$ is unbounded, by lemma 5.4.3.

By continuity,

$$\begin{aligned}
\lim_{p\to\infty} b_{p,n-s}(t_\infty) &= \lim_{p\to\infty}\frac{1}{s!}\left(\frac{\partial}{\partial x}\right)^{s}\Bigg|_{x=x_1(t_p,y_p)} P(t_\infty,y_p)(x)\\
&= \frac{1}{s!}\left(\frac{\partial}{\partial x}\right)^{s}\Bigg|_{x=x_1(t_\infty,y_\infty)} P(t_\infty,y_\infty)(x)\\
&\neq 0,
\end{aligned}$$

since s is the multiplicity of $x_1(t_\infty,y_\infty)$. Consequently, there exists a subsequence such that for a suitable neighborhood U of t_∞

$$\inf\{|b_{p,n-s}(t)| : t\in U, (t_p,y_p) \text{ in the subsequence}\} > 0.$$

By dilating the t-axis and denoting the subsequence again by $(t_p,y_p)_p$, we can assume without loss of generality that

$$\inf\{|b_{p,n-s}(t)| : |t-t_\infty|\le 1, p\in\mathbb{N}\} > 0$$

which implies that $b_{p,n-s}(t)\neq 0$, for all $|t-t_\infty|\le 1$ and all $p\in\mathbb{N}$.

Next we observe that

$$b_{p,n-q}(t_p) = \frac{1}{q!}\left(\frac{\partial}{\partial x}\right)^q\bigg|_{x=x_1(t_p,y_p)} P(t_p,y_p)(x) \neq 0$$

and, for $1 \leq j \leq q$,

$$b_{p,n-q+j}(t_p) = \frac{1}{(q-j)!}\left(\frac{\partial}{\partial x}\right)^{q-j}\Bigg|_{x=x_1(t_p,y_p)} P(t_p,y_p)(x) = 0,$$

since the multiplicity of $x_1(t_p, y_p)$ is q.

Let us consider

$$\left(\frac{\partial}{\partial \tilde{x}}\right)^{q-j} Q_p(t)(\tilde{x}) = \sum_{i=0}^{n-q+j} \frac{(n-i)!}{(n-i-q+j)!} b_{p,i}(t)\tilde{x}^{n-i-q+j},$$

for $j = 1, 2$ and $j \leq q$. We want to apply lemma 5.5.1 to this polynomial, if $j = 1$, and lemma 5.5.12, if $j = 2$, respectively. The correspondence between the present and the former used notation (up to unimportant constant factors) is the following:

$$\begin{array}{ccc} A_0 & \longleftrightarrow & b_{p,n-s} \\ A_1 & \longleftrightarrow & b_{p,n-s+1} \\ \vdots & \vdots & \vdots \\ A_{s-q+j} & \longleftrightarrow & b_{p,n-q+j}. \end{array}$$

Still to check are the differentiability conditions in lemma 5.5.1 and lemma 5.5.12. All $b_{p,0}, \ldots, b_{p,n}$ are of class C^k, by assumption. Hence, to show is that $s-q+j \leq k$. But this is obvious, since $j \leq q$.

Thus, the application of lemma 5.5.1 and lemma 5.5.12 gives

$$\left|\frac{b^{(j)}_{p,n-q+j}(t_p)}{b_{p,n-q}(t_p)}\right| \leq C_1 \qquad (j = 1, 2),$$

where C_1 is a constant not depending on p, since $(t_p, y_p)_p \subseteq K$. Therefore,

$$\sup_{p\in\mathbb{N}} \left|\frac{b^{(j)}_{p,n-q+j}(t_p)}{b_{p,n-q}(t_p)}\right| < \infty \qquad (j = 1, 2). \tag{5.37}$$

By the lemma below, (5.37) holds for all $1 \leq j \leq q$, and, hence, all coefficients of $(b_{p,n-q}(t_p))^{-1}T_p(x)$ are bounded in p. □

LEMMA. *Consider a sequence $(P_m)_{m\in\mathbb{N}}$ of hyperbolic polynomials*

$$P_m(x) = x^n - a_{m,1}x^{n-1} + \cdots + (-1)^n a_{m,n}.$$

If the first two coefficients $a_{m,1}$ and $a_{m,2}$ are bounded, then all other coefficients $a_{m,i}$ $(3 \leq i \leq n)$ are bounded as well.

PROOF. Let $x_{m,1}, \ldots, x_{m,n}$ denote the roots of P_m and let $s_{m,2}$ be the second Newton polynomial in the roots. By (3.1), we have

$$x^2_{m,1} + \cdots + x^2_{m,n} = s_{m,2} = a^2_{m,1} - 2a_{m,2}.$$

If $a_{m,1}$ and $a_{m,2}$ are bounded, all roots $x_{m,1}, \ldots, x_{m,n}$ are bounded, and, via Vieta's formulas, also all other coefficients $a_{m,i}$ $(3 \leq i \leq n)$ are bounded. □

5.5.14. COROLLARY. *Suppose that the polynomial*

$$P(t,y)(x) = x^n - a_1(t,y)x^{n-1} + \cdots + (-1)^n a_n(t,y)$$

is hyperbolic for any $(t,y) \in (-1,1)^m \times \mathcal{M}$, where $\mathcal{M}$ is a compact Hausdorff topological space, and the multiplicity of its roots does not exceed k. Furthermore, *suppose that all partial derivatives* $\partial_t^\alpha a_j(t,y)$ $(0 \le |\alpha| \le k; 1 \le j \le n)$ *are continuous functions on* $(-1,1)^m \times \mathcal{M}$*, where* $\partial_t^\alpha = \left(\frac{\partial}{\partial t_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial t_m}\right)^{\alpha_m}$*. Then there is a continuous parameterization of the roots of* P *which is locally Lipschitz in* t*.*

PROOF. This follows from theorem 5.5.13 and the fact that a mapping between Banach spaces is locally Lipschitz if and only if it is locally Lipschitz along each C^∞ curve (cf. [**KM97**, 12.7]). □

CHAPTER 6

Wakabayashi's approach

In this chapter we shall present a more general version of Bronshein's theorem 5.5.13 due to Wakabayashi who published it in 1986, see [**Wak86**]. Wakabayashi's approach is shorter and more conceptual than Bronshtein's.

6.1. Preliminaries

Throughout this section let $P(x) = x^n + \sum_{j=1}^n a_j x^{n-j}$ be a monic polynomial, with coefficients in $\mathbb{C}$ and viewed as function on $\mathbb{C}$, if not stated otherwise.

We shall use the splitting operator $P \mapsto P + sP'$ $(s \in \mathbb{C})$ that reduces the multiplicity of the multiple roots of P. Let us observe at first that the hyperbolicity of polynomials remains invariant under this operator.

6.1.1. The operator $P \mapsto P + sP'$ preserves hyperbolicity.

LEMMA. *If* $P(x) \neq 0$ *for* $\mathrm{Im}(x) < 0$*, then* $\left(1 + s\frac{d}{dx}\right) P(x) \neq 0$ *for* $\mathrm{Im}(x) < 0$ *and* $\mathrm{Im}(s) \leq 0$.

PROOF. Let $\alpha_1, \ldots, \alpha_n$ be the roots of P such that $P(x) = \prod_{j=1}^n (x - \alpha_j)$. Then, by assumption, $\mathrm{Im}(\alpha_j) \geq 0$, for $1 \leq j \leq n$. Consider

$$\left(1 + s\frac{d}{dx}\right) P(x) = P(x)(1 + s\sum_{j=1}^n (x - \alpha_j)^{-1}).$$

Now, suppose that $\mathrm{Im}(x) < 0$. Then, $x - \alpha_j \neq 0$, and

$$(x - \alpha_j)^{-1} = (\mathrm{Re}(x - \alpha_j) + i\mathrm{Im}(x - \alpha_j))^{-1} = \frac{\mathrm{Re}(x - \alpha_j) - i\mathrm{Im}(x - \alpha_j)}{|x - \alpha_j|^2},$$

so

$$\mathrm{Im}((x - \alpha_j)^{-1}) = -\frac{\mathrm{Im}(x - \alpha_j)}{|x - \alpha_j|^2} = \frac{\mathrm{Im}(\alpha_j) - \mathrm{Im}(x)}{|x - \alpha_j|^2} > 0.$$

Note that the statement of the lemma is trivial, if $s = 0$.

For contradiction, let us assume that, for $\mathrm{Im}(x) < 0$, $s \neq 0$, and $\mathrm{Im}(s) \leq 0$, we have $\left(1 + s\frac{d}{dx}\right) P(x) = 0$. Then, $1 + s\sum_{j=1}^n (x - \alpha_j)^{-1} = 0$. But this means that

$$\mathrm{Re}(s\sum_{j=1}^n (x - \alpha_j)^{-1}) = -1 \quad \text{and} \quad \mathrm{Im}(s\sum_{j=1}^n (x - \alpha_j)^{-1}) = 0. \tag{6.1}$$

To shorten notation let us write $u = \sum_{j=1}^n (x - \alpha_j)^{-1}$, and we have $\mathrm{Im}(u) = \sum_{j=1}^n \mathrm{Im}((x - \alpha_j)^{-1}) > 0$. Then, the equations in (6.1) take the following form

$$\begin{aligned} -1 &= \mathrm{Re}(su) = \mathrm{Re}(s)\mathrm{Re}(u) - \mathrm{Im}(s)\mathrm{Im}(u) \\ 0 &= \mathrm{Im}(su) = \mathrm{Re}(s)\mathrm{Im}(u) + \mathrm{Im}(s)\mathrm{Re}(u). \end{aligned}$$

The second equation implies that $\mathrm{Re}(s)$ and $\mathrm{Re}(u)$ have the same sign, whence $\mathrm{Re}(s)\mathrm{Re}(u) - \mathrm{Im}(s)\mathrm{Im}(u) \geq 0$, contradicting the first equation. This completes the proof of the lemma. $\square$

Corollary. *Clearly, the statement of the previous lemma remains true, if all order-relations are reversed. Consequently, if P is hyperbolic, then so is $\left(1+s\frac{d}{dx}\right)P$ (and iterations), for $s \in \mathbb{R}$.*

6.1.2. The operator $P \mapsto P+sP'$ reduces the multiplicity. The following lemma shows that, indeed, the operator $P \mapsto P+sP'$ reduces the multiplicity of the roots (see (6.2)), and it gives an estimate for the deviation the roots are subjected to under this operator (see (6.3)).

Lemma. *Let $P(x) = \prod_{j=1}^n (x-\alpha_j^0)$ be a hyperbolic polynomial with roots $\alpha_1^0 \leq \alpha_2^0 \leq \cdots \leq \alpha_n^0$. For $s \in \mathbb{R}$ let us consider the hyperbolic polynomial*

$$\left(1+s\frac{d}{dx}\right)^{n-1} P(x) = \prod_{j=1}^n (x-\alpha_j(s)),$$

where $\alpha_1(s) \leq \alpha_2(s) \leq \cdots \leq \alpha_n(s)$ and $\alpha_j(0) = \alpha_j^0$, for $1 \leq j \leq n$. Then, there exist positive constants $C_1(n)$ and $C_2(n)$, depending only on n, such that

$$\alpha_j(s) - \alpha_{j-1}(s) \geq C_1(n)|s| \quad \textit{for } s \in \mathbb{R} \textit{ and } 2 \leq j \leq n \tag{6.2}$$

and

$$0 < \pm(\alpha_j^0 - \alpha_j(s)) \leq C_2(n)|s| \quad \textit{for } \pm s > 0 \textit{ and } 1 \leq j \leq n. \tag{6.3}$$

Proof. First of all note that, for $s = 0$, (6.2) is trivial. Consider the case where $s > 0$. We make induction on the number how often the operator $1+s\frac{d}{dx}$ is applied to P: Assume that, for a fixed $1 \leq l \leq n-1$, there is a positive constant $C_1(l)$ such that

$$\alpha_j^l(s) - \alpha_{j-1}^l(s) \geq C_1(l)s \quad \text{for } s > 0 \text{ and } 2 \leq j \leq l, \tag{6.4}$$

where $\left(1+s\frac{d}{dx}\right)^{l-1} P(x) = \prod_{j=1}^n (x-\alpha_j^l(s))$ and $\alpha_1^l(s) \leq \alpha_2^l(s) \leq \cdots \leq \alpha_n^l(s)$. Note that (6.4)) is trivially satisfied, if $l = 1$. Put

$$\begin{aligned}
f(x,s) &= \frac{\left(1+s\frac{d}{dx}\right)^l P(x)}{\left(1+s\frac{d}{dx}\right)^{l-1} P(x)} = \frac{\left(1+s\frac{d}{dx}\right)\left(1+s\frac{d}{dx}\right)^{l-1} P(x)}{\left(1+s\frac{d}{dx}\right)^{l-1} P(x)} \\
&= 1 + s\frac{\frac{d}{dx}\left(1+s\frac{d}{dx}\right)^{l-1} P(x)}{\left(1+s\frac{d}{dx}\right)^{l-1} P(x)} \\
&= 1 + s\frac{\sum_{j=1}^n (x-\alpha_1^l(s))\cdots\widehat{(x-\alpha_j^l(s))}\cdots(x-\alpha_n^l(s))}{\prod_{j=1}^n (x-\alpha_j^l(s))} \\
&= 1 + s\sum_{j=1}^n (x-\alpha_j^l(s))^{-1},
\end{aligned}$$

where '$\widehat{\ }$' stands for omission. Since $s > 0$, we find, for $1 \leq h \leq n$ and $\alpha_{h-1}^l(s) < x < \alpha_h^l(s)$,

$$1 + sn(x-\alpha_1^l(s))^{-1} < f(x,s) < 1 + s(x-\alpha_1^l(s))^{-1}, \quad \text{when } h = 1 \tag{6.5}$$

and

$$\begin{aligned}
&f(x,s) > 1 + s(x-\alpha_{h-1}^l(s))^{-1} + s(n-h+1)(x-\alpha_h^l(s))^{-1} \quad \text{and} \\
&f(x,s) < A_h + s(x-\alpha_{h-1}^l(s))^{-1} + s(x-\alpha_h^l(s))^{-1}, \quad \text{when } 2 \leq h \leq n,
\end{aligned} \tag{6.6}$$

where we set $\alpha_0^l(s) = -\infty$, $A_2 = 1$, and $A_h = 1 + s(h-2)(\alpha_{h-1}^l(s) - \alpha_{h-2}^l(s))^{-1}$, if $3 \le h \le n$. In fact,

$$\begin{aligned} f(x,s) &= 1 + s\sum_{j=1}^{h-2} \underbrace{(x-\alpha_j^l(s))^{-1}}_{>0} + s(x-\alpha_{h-1}^l(s))^{-1} + s\sum_{j=h}^{n} \underbrace{(x-\alpha_j^l(s))^{-1}}_{\ge (x-\alpha_h^l(s))^{-1}} \\ &> 1 + s(x-\alpha_{h-1}^l(s))^{-1} + s(n-h+1)(x-\alpha_h^l(s))^{-1} \end{aligned}$$

and

$$\begin{aligned} f(x,s) = 1 &+ s\sum_{j=1}^{h-2} \underbrace{(x-\alpha_j^l(s))^{-1}}_{<(\alpha_{h-1}^l(s)-\alpha_{h-2}^l(s))^{-1}} + s(x-\alpha_{h-1}^l(s))^{-1} + s(x-\alpha_h^l(s))^{-1} \\ &+ s\sum_{j=h+1}^{n} \underbrace{(x-\alpha_j^l(s))^{-1}}_{<0} \\ &< A_h + s(x-\alpha_{h-1}^l(s))^{-1} + s(x-\alpha_h^l(s))^{-1}. \end{aligned}$$

We assert that, for $1 \le h \le n$ and $\alpha_{h-1}^l(s) < \alpha_h^l(s)$, this yields

$$\begin{cases} \alpha_{h-1}^l(s) < \alpha_h^{l+1}(s) < \alpha_h^l(s), \\ \alpha_1^l(s) - sn < \alpha_1^{l+1}(s) < \alpha_1^l(s) - s \quad \text{when } h = 1, \\ \alpha_h^l(s) - \frac{1}{2}\left(X_h + s(n-h+2) - [(X_h - s(n-h+2))^2 + 4sX_h]^{\frac{1}{2}}\right) \\ \quad < \alpha_h^{l+1}(s) < \alpha_h^l(s) - \frac{1}{2}F(X_h, \frac{2s}{A_h}) \quad \text{when } 2 \le h \le n, \end{cases} \tag{6.7}$$

with $X_h = \alpha_h^l(s) - \alpha_{h-1}^l(s)$ and $F(u,v) = u + v - (u^2+v^2)^{\frac{1}{2}}$. To prove the inequalities in the first row of (6.7) we introduce the following notation (for fixed $s > 0$)

$$R(x) = \left(1 + s\frac{d}{dx}\right)^l P(x) = \left(1 + s\frac{d}{dx}\right)\left(1 + s\frac{d}{dx}\right)^{l-1} P(x) = Q(x) + sQ'(x),$$

and we observe that, at roots of Q, the polynomials R and Q' have the same sign. Now, let us apply this to $\alpha_1^l(s)$ which is the smallest root of Q. Therefore, we find that $\alpha_1^{l+1}(s) \le \alpha_1^l(s)$, since $\alpha_1^{l+1}(s)$ is the smallest root of R and since R and Q have the same asymptotic behavior for $x \to -\infty$. If we consider two consecutive roots of Q, say $\alpha_h^l(s)$ and $\alpha_{h+1}^l(s)$ with $1 \le h \le n-1$, then either they coincide or Q' takes different signs at them. In both cases there has to be a root of R between them. In particular, if $\alpha_h^l(s) = \cdots = \alpha_{h+k}^l(s)$ is a $k+1$-fold root of Q, then it has to coincide with at least k roots of R. We have shown that in each of the intervals (maybe consisting of one single point) $(-\infty, \alpha_1^l(s)], [\alpha_1^l(s), \alpha_2^l(s)], \ldots, [\alpha_{n-1}^l(s), \alpha_n^l(s)]$ lies a root of R, thus, R having the same degree as Q,

$$\alpha_1^{l+1}(s) \le \alpha_1^l(s) \le \alpha_2^{l+1}(s) \le \cdots \le \alpha_n^{l+1}(s) \le \alpha_n^l(s).$$

If in particular $\alpha_{h-1}^l(s) < \alpha_h^l(s)$, then we obtain $\alpha_{h-1}^l(s) < \alpha_h^{l+1}(s) < \alpha_h^l(s)$. For, if $\alpha_{h-1}^l(s) = \alpha_h^{l+1}(s)$, then $0 = R(\alpha_h^{l+1}(s)) = Q(\alpha_{h-1}^l(s)) + sQ'(\alpha_{h-1}^l(s)) = sQ'(\alpha_{h-1}^l(s))$, whence $\alpha_{h-1}^l(s)$ is a multiple root. By assumption, $\alpha_{h-1}^l(s)$ cannot equal $\alpha_h^l(s)$, so $\alpha_{h-2}^l(s) = \alpha_{h-1}^l(s)$. But then $\alpha_{h-1}^{l+1}(s) = \alpha_h^{l+1}(s)$ is a multiple root of R, that means $0 = R'(\alpha_h^{l+1}(s)) = Q'(\alpha_{h-1}^l(s)) + sQ''(\alpha_{h-1}^l(s)) = sQ''(\alpha_{h-1}^l(s))$. Therefore, $\alpha_{h-1}^l(s)$ has to be an at least 3-fold root. By continuing this procedure we finally see that $\alpha_{h-1}^l(s)$ must be an h-fold root of Q, whence $\alpha_{h-1}^l(s) = \alpha_h^l(s)$, in contradiction to the assumption. This shows the first row of (6.7), since the second inequality is analogous.

The second row of (6.7) is obtained by applying (6.5) with $x = \alpha_1^{l+1}(s)$, note that $f(\alpha_1^{l+1}(s), s) = 0$. The third one can be derived in the same way from (6.6), by putting $x = \alpha_h^{l+1}(s)$. Thus, the assertion is shown.

Since $(X_h - (n-h+2)s)^2 + 4sX_h = (X_h - (n-h)s)^2 + 4(n-h+1)s^2 \geq (X_h - (n-h)s)^2$, (6.7) yields

$$0 < \alpha_h^l(s) - \alpha_h^{l+1}(s) < (n-h+1)s, \text{ for } 1 \leq h \leq n. \tag{6.8}$$

Moreover, (6.7) implies that

$$\begin{aligned} \alpha_{h+1}^{l+1}(s) - \alpha_h^{l+1}(s) &= \underbrace{\alpha_{h+1}^{l+1}(s) - \alpha_h^l(s)}_{\geq 0} + \alpha_h^l(s) - \alpha_h^{l+1}(s) \\ &\geq \begin{cases} s & \text{if } h = 1, \\ \frac{1}{2}F(X_h, \frac{2s}{A_h}) & \text{if } 2 \leq h \leq n. \end{cases} \end{aligned}$$

By induction hypothesis (6.4), we have $X_h = \alpha_h^l(s) - \alpha_{h-1}^l(s) \geq C_1(l)s$ and $A_h = 1 + s(h-2)(\alpha_{h-1}^l(s) - \alpha_{h-2}^l(s))^{-1} \leq 1 + (h-2)C_1(l)^{-1}$, for $h \leq l$. This and the fact that F is positively homogeneous and satisfies $F(u_1, v_1) \geq F(u_2, v_2)$, if $u_1 \geq u_2 \geq 0$ and $v_1 \geq v_2 \geq 0$, imply

$$\alpha_{h+1}^{l+1}(s) - \alpha_h^{l+1}(s) \geq \begin{cases} s & \text{if } h = 1, \\ \frac{s}{2}F(C_1(l), \frac{2C_1(l)}{h-2+C_1(l)}) & \text{if } 2 \leq h \leq l. \end{cases}$$

Hence, we have shown that (6.4) is also valid, if we replace l by $l+1$, where $C_1(l+1) = \min\left\{1, \frac{1}{2}F(C_1(l), \frac{2C_1(l)}{h-2+C_1(l)})\right\}$. This proves (6.2), for $s > 0$.

To get (6.3), for $s > 0$, note that, by definition, $\alpha_j^0 = \alpha_j^1(s)$ and $\alpha_j(s) = \alpha_j^n(s)$. Then, (6.8) yields

$$0 < \alpha_j^0 - \alpha_j(s) = \sum_{l=1}^{n-1}(\alpha_j^l(s) - \alpha_j^{l+1}(s)) \leq (n-1)(n-j+1)s \leq n(n-1)s,$$

for all $1 \leq j \leq n$. This completes the proof, when $s > 0$. Similarly, one can prove the lemma, when $s < 0$. □

6.1.3. Lemma. *Consider $P(x) = x^n + \sum_{j=1}^n a_j x^{n-j}$ and $Q(x) = \sum_{j=1}^n b_j x^{n-j}$, and write $P(x) + Q(x) = \prod_{j=1}^n (x - \alpha_j(b_1, \ldots, b_n))$. Then there exists a positive constant $C(n)$, depending only on n, such that*

$$|\alpha_j(b_1, \ldots, b_n) - \alpha_j^0| \leq C(n) \max_{1 \leq k \leq n} \left(|b_k|^{\frac{1}{k}} + |b_k|^{\frac{1}{n}} |\alpha_j^0|^{1-\frac{k}{n}}\right), \quad 1 \leq j \leq n, \tag{6.9}$$

where $\alpha_j^0 = \alpha_j(0, \ldots, 0)$, $1 \leq j \leq n$, are the roots of P.

Proof. We shall prove that (6.9) holds for $j = 1$. The remaining cases $2 \leq j \leq n$ are identical. There is a integer k_0 with $1 \leq k_0 \leq n$ such that none of $\alpha_2^0, \ldots, \alpha_n^0$ lies in $U := \{z \in \mathbb{C} : (k_0 - 1)A \leq |z - \alpha_1^0| < k_0 A\}$, where $A > 0$ is determined later. Geometrically speaking, U is the region between two circles both with center α_1^0 (or, if $k_0 = 1$, the inner circle shrinks to the point α_1^0) in the complex plane. Our goal is to prove that on the middle-circle of U, namely $|z - \alpha_1^0| = (k_0 - 2^{-1})A$, we have $|P(z)| > |Q(z)|$, in order to apply Rouché's theorem 2.2.2. If $|z - \alpha_1^0| = (k_0 - 2^{-1})A$, then we have $|P(z)| = |\prod_{j=1}^n (z - \alpha_j^0)| = \prod_{j=1}^n |z - \alpha_j^0| \geq \left(\frac{A}{2}\right)^n$ and, of course, $|Q(z)| \leq \sum_{j=1}^n |b_j||z|^{n-j}$. Consequently, we find

$$|P(z)| - |Q(z)| \geq \left(\frac{A}{2}\right)^n - \sum_{j=1}^n |b_j||z|^{n-j},$$

if $|z - \alpha_1^0| = (k_0 - 2^{-1})A$.

We assert that there is a $C'(n) > 0$, depending only on n, such that

$$\left(\frac{A}{2}\right)^n > n|b_i|\left(|\alpha_1^0| + (k_0 - 2^{-1})A\right)^{n-i},$$

for $1 \le i \le n$, $A \ge C'(n)\left(|b_i|^{\frac{1}{i}} + |b_i|^{\frac{1}{n}}|\alpha_1^0|^{1-\frac{i}{n}}\right)$, and $b_i \neq 0$. Indeed, the required estimate is a polynomial inequality in A, whence the leading term A^n dominates the rest whenever A is large enough. We can achieve it by a suitable choice of $C'(n)$, and the assertion follows.

Then, we find

$$|P(z)| - |Q(z)| > \sum_{j=1}^{n} |b_j|\left(|\alpha_1^0| + (k_0 - 2^{-1})A\right)^{n-j} - \sum_{j=1}^{n} |b_j||z|^{n-j}.$$

Since $|\alpha_1^0| + (k_0 - 2^{-1})A \ge |z|$, this implies that on the circle $|z - \alpha_1^0| = (k_0 - 2^{-1})A$ the inequality $|P(z)| > |Q(z)|$ is satisfied. Applying Rouché's theorem 2.2.2, we obtain

$$|\alpha_1(b_1, \ldots, b_n) - \alpha_1^0| \le (k_0 - 2^{-1})A \le (n - 2^{-1})A,$$

and, if we put $A := C'(n)\max_{1\le k\le n}\left(|b_k|^{\frac{1}{k}} + |b_k|^{\frac{1}{n}}|\alpha_1^0|^{1-\frac{k}{n}}\right)$, then the lemma is proved for $j = 1$. □

6.1.4. Lemma. *Let M be an arc-wise connected subset of $\mathbb{R}^n$, U a Hausdorff topological space, and $S = \{s \in \mathbb{C} : |s| \le s_0$ and $\mathrm{Im}(s) \le 0\}$, with $s_0 \in \mathbb{R}_+$. Suppose $f : S \times M \times U \to \mathbb{C}$ is a continuous function that satisfies the following conditions:*

(i) *$f(s, w, u)$ is holomorphic in s, if $\mathrm{Im}(s) < 0$.*
(ii) *There is a dense subset U' of U such that $f(s, w, u) \neq 0$, for $s \in S \cap \mathbb{R}$, $w \in M$, and $u \in U'$.*
(iii) *$f(s, w, u) \neq 0$, if $|s| = s_0$.*
(iv) *There is a $w_0 \in M$ such that $f(s, w_0, u) \neq 0$, if $\mathrm{Im}(s) < 0$.*

Then, $f(s, w, u) \neq 0$, if $\mathrm{Im}(s) < 0$.

Proof. For contradiction, assume that there is an element $(s_1, w_1, u_1) \in S \times M \times U$ such that $\mathrm{Im}(s_1) < 0$ and $f(s_1, w_1, u_1) = 0$.

First we assert that we can suppose without loss of generality that $u_1 \in U'$. Suppose $u_1 \notin U'$. Condition (iii) implies that $s \mapsto f(s, w_1, u_1)$ cannot vanish identically on S. So s_1 is an isolated root of $s \mapsto f(s, w_1, u_1)$, and we may find a small circle γ in $\{s \in \mathbb{C} : \mathrm{Im}(s) < 0\}$ centered at s_1 such that s_1 is the only zero of $s \mapsto f(s, w_1, u_1)$ inside and on γ. Since U' is dense in U and by continuity, there exists a $u' \in U'$ sufficiently close to u_1 such that $|f(s, w_1, u_1)| > |f(s, w_1, u') - f(s, w_1, u_1)|$ holds for all $s \in \gamma$. Application of Rouché's theorem 2.2.2 yields that there has to be a s' with $\mathrm{Im}(s') < 0$ such that $f(s', w_1, u') = 0$. This shows the assertion.

Since M is arc-wise connected, we can find a curve $c_M : [0, 1] \to M$ with $c_M(0) = w_1$ and $c_M(1) = w_0$. Then, by the conditions (i) - (iii), we can apply theorem 2.2.3 which implies that there is a curve $c_S : [0, 1] \to S$ with $c_S(0) = s_1$ such that $f(c_S(t), c_M(t), u_1) = 0$ for $t \in [0, 1]$. Observe that the entire curve c_S lies in the interior of S, by (iii) and since we have $u_1 \in U'$. Consequently, we have $f(c_S(1), w_0, u_1) = 0$, where $\mathrm{Im}(c_S(1)) < 0$, contradicting condition (iv). This proves the lemma. □

6.2. Microhyperbolicity

We introduce the notion of microhyperbolicity and discuss some properties which will be needed in the next section. The following considerations are based on [**Hör63**] and [**Hör83a**].

6.2.1. Definition. A real analytic function F on an open set $U \subseteq \mathbb{R}^n$ is called *microhyperbolic with respect to* $\Theta \in \mathbb{R}^n$, if there is a positive continuous function $x \mapsto t(x)$ on U such that $F(x+it\Theta) \neq 0$, for $0 < t < t(x)$ and $x \in U$.

In the following discussion of the local properties of F we may shrink U so that the function $x \mapsto t(x)$ is bounded from below on U by a positive constant and then replace Θ by a multiple to achieve that

$$F(x+it\Theta) \neq 0, \quad \text{if } 0 < t \leq 1 \text{ and } x \in U. \tag{6.10}$$

6.2.2. Lemma. *If F satisfies* (6.10) *and $F(x_0+t\Theta)$ has a zero of multiplicity m exactly when $t=0$, where $x_0 \in U$, then*

$$F(x) = F_0(x) + O(|x-x_0|^{m+1}), \quad \textit{for } x \in U,$$

where F_0 is a homogeneous polynomial of degree m with

$$F_0(\Theta) \neq 0 \quad \textit{and} \quad F_0(x+it\Theta) \neq 0, \ \textit{if } t \in \mathbb{R}\backslash\{0\} \textit{ and } x \in \mathbb{R}^n. \tag{6.11}$$

Proof. To simplify notation we assume without loss of generality that $0 \in U$ and $x_0 = 0$. Let $y \in \mathbb{R}^n$ be a fixed vector and set $g(t,s) := F(t\Theta + sy)$ for real s. Then $g(t,0) = F(t\Theta) = ct^m + O(t^{m+1})$ with $c \neq 0$, since $F(t\Theta)$ vanishes of order m exactly at $t = 0$.

We claim that $g(t,s) = O(|t|+|s|)^m$ at $(0,0)$. Suppose this is not true. Consider the largest $\lambda \in \mathbb{R}$ such that $g(t,s) = O(|t|+|s|^\lambda)^m$ at $(0,0)$. Then, we find $\frac{1}{m} \leq \lambda < 1$, since $g(0,s) = F(sy)$ vanishes at $s=0$. Moreover, F and thus also g being real analytic, λ has to be a rational number. Write $\lambda = \frac{p}{q}$, where $1 \leq p < q$ are relatively prime integers. Let us consider the limits

$$g_0^{\pm}(w) := \lim_{s \to \pm 0} \frac{g(w|s|^\lambda, s)}{|s|^{m\lambda}},$$

where $w \in \mathbb{C}$. If at^js^k is a term in the expansion of $g(t,s)$ with $j + \frac{k}{\lambda} = m$, then q divides $m-j$, because then we have $p(m-j) = kq$, and p and q are relatively prime. In the expansion of $|s|^{-m\lambda}g(w|s|^\lambda, s)$ only terms of the form $|s|^{-m\lambda}aw^j|s|^{\lambda j}s^k$ with $j + \frac{k}{\lambda} = m$ survive as $|s| \to 0$, which follows from $g(t,s) = O(|t|+|s|^\lambda)^m$ at $(0,0)$. Consequently, $k = \lambda(m-j) = \frac{p}{q}(m-j) = pl$ with $l \in \mathbb{N}_0$ which implies $j = m - \frac{k}{\lambda} = m - ql$, and so we obtain

$$|s|^{-m\lambda}aw^j|s|^{\lambda j}s^k = |s|^{-m\lambda}aw^j|s|^{\lambda j}|s|^k \operatorname{sgn}(s)^k = aw^j \operatorname{sgn}(s)^{pl} = aw^{m-ql}\operatorname{sgn}(s)^{pl}.$$

Therefore, we have

$$\begin{aligned} g_0^{\pm}(w) &= cw^m + (\pm 1)^p c_1 w^{m-q} + (\pm 1)^{2p} c_2 w^{m-2q} + \cdots + (\pm 1)^{dp} c_d w^r \\ &= w^r \left(cw^{m-r} + (\pm 1)^p c_1 w^{m-q-r} + \cdots + (\pm 1)^{dp} c_d \right), \end{aligned}$$

with $c \neq 0$ and not all $c_j = 0$, where $m = dq + r$ is the division of m by q with remainder r. The second factor on the right-hand side of the above equation is a polynomial in $w^q =: z$ of degree d; let us denote it by $h_0^{\pm}(z)$. We can find a non-zero root z_0 of $cz^d + c_1 z^{d-1} + \cdots + c_d = 0$, since c and at least one of the c_j do not vanish. So we have

$$\begin{aligned} h_0^{\pm}((\pm 1)^p z_0) &= c(\pm 1)^{dp} z_0^d + (\pm 1)^p c_1 (\pm 1)^{(d-1)p} z_0^{d-1} + \cdots + (\pm 1)^{dp} c_d \\ &= (\pm 1)^{dp} \left(cz_0^d + c_1 z_0^{d-1} + \cdots + c_d \right) = 0. \end{aligned}$$

Thus, $g_0^{\pm}(w) = 0$, if $w^q = (\pm 1)^p z_0$. All such w cannot lie in a half-plane, unless $q = 2$ and p is even which has been excluded by requiring $1 \leq p < q$. On the other hand the roots of $g_0^{\pm}(w) = 0$ satisfy $\operatorname{Im}(w) \leq 0$, for, if

$$g(w|s|^\lambda, s) = F(\operatorname{Re}(w)|s|^\lambda\Theta + sy + i\operatorname{Im}(w)|s|^\lambda\Theta) = 0$$

and s is sufficiently small, then $\mathrm{Im}(w)$ cannot be positive, by (6.10). So the assumption $\lambda \neq 1$ led to a contradiction and, thus, the assertion is established.

Since $y \in \mathbb{R}^n$ was arbitrary, we conclude that $F(x) = O(|x|^m)$ as $x \to 0$ (set $y = \frac{1}{|x|}x - \Theta$). Now define

$$F_0(x) := \lim_{\epsilon \searrow 0} \frac{F(\epsilon x)}{\epsilon^m}.$$

Then F_0 is evidently a homogeneous polynomial of degree m. The first part of (6.11), namely $F_0(\Theta) \neq 0$, is a direct consequence of the definition of F_0 and the assumption that $F(\epsilon\Theta)$ vanishes of order m exactly at $\epsilon = 0$. Moreover, considering $F(\epsilon(x + w\Theta)) = F(\epsilon x + \epsilon\mathrm{Re}(w)\Theta + i\epsilon\mathrm{Im}(w)\Theta)$ for small $\epsilon > 0$ in addition with (6.10), yields $F_0(x + w\Theta) \neq 0$, if $x \in \mathbb{R}^n$ and $\mathrm{Im}(w) > 0$. Hence $F_0(x + w\Theta) = (-1)^m F_0(-x - w\Theta) \neq 0$, if $x \in \mathbb{R}^n$, and $\mathrm{Im}(w) < 0$. This shows the second part of (6.11) and completes the proof. □

The polynomial F_0 appearing in the previous lemma is often referred to as the *localization polynomial.*

6.3. Hölder continuity of the roots

The following theorem provides a variant of Bronshtein's theorem 5.5.13.

6.3.1. THEOREM. *Consider an open convex subset $T \subseteq \mathbb{R}^m$ and a compact Hausdorff topological space $\mathcal{Y}$. Let $P(t, y)(x) = x^n + \sum_{j=1}^n a_j(t, y)x^{n-j}$ be a monic polynomial, where the coefficients $a_1, \dots, a_n$ are real-valued functions defined for $t = (t_1, \dots, t_m) \in T$ and $y \in \mathcal{Y}$. Assume that $P(t, y)$ is hyperbolic for all $(t, y) \in T \times \mathcal{Y}$. Moreover, we suppose that all partial derivatives $\partial_t^\alpha a_j(t, y)$ ($|\alpha| \leq k, 1 \leq j \leq n$) are continuous on $T \times \mathcal{Y}$ and that there exist constants $C > 0$ and $0 < \delta \leq 1$ such that*

$$|\partial_t^\alpha a_j(t, y) - \partial_t^\alpha a_j(t', y)| \leq C|t - t'|^\delta, \tag{6.12}$$

for $|\alpha| = k$, $t, t' \in T$, and $y \in \mathcal{Y}$, where k is a non-negative integer and $\partial_t^\alpha = \left(\frac{\partial}{\partial t_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial t_m}\right)^{\alpha_m}$. Then, for any open relative-compact subset $U \subseteq T$, there is a $C_U > 0$ such that the ordered roots $\lambda_1(t, y) \leq \dots \leq \lambda_n(t, y)$ of $P(t, y)$ satisfy

$$|\lambda_j(t, y) - \lambda_j(t', y)| \leq C_U|t - t'|^r, \tag{6.13}$$

for $1 \leq j \leq n$, $t, t' \in U$, and $y \in \mathcal{Y}$, where $r = \min\left\{1, \frac{k+\delta}{n}\right\}$.

The proof of theorem 6.3.1 will run from 6.3.2 till 6.3.12.

6.3.2. We set

$$\tilde{P}(t, y, z)(x) = \left(1 + z^r \frac{\partial}{\partial x}\right)^{n-1} P(t, y)(x),$$

for $z \in \mathbb{C}$ with $\mathrm{Im}(z) \leq 0$, where we demand $z^r = |z|^r e^{-ir\pi}$ if $z \leq 0$. Corollary 6.1.1 implies that $\tilde{P}(t, y, z)$ is hyperbolic, if $(t, y) \in T \times \mathcal{Y}$ and $z^r \in \mathbb{R}$. The condition $z^r \in \mathbb{R}$ is equivalent to either $z \in \mathbb{R}$, if $r = 1$, or $z \in \mathbb{R}$ and non-negative, if $r = \frac{k+\delta}{n} < 1$. Moreover, as a consequence of lemma 6.1.2, if $z \geq 0$, or $z \in \mathbb{R}$ and $r = 1$, then we find positive constants $C_1(n)$ and $C_2(n)$ such that

$$\tilde{\lambda}_j(t, y, z) - \tilde{\lambda}_{j-1}(t, y, z) \geq C_1(n)|z|^r, \quad 2 \leq j \leq n \tag{6.14}$$

and

$$|\lambda_j(t, y) - \tilde{\lambda}_j(t, y, z)| \leq C_2(n)|z|^r, \quad 1 \leq j \leq n, \tag{6.15}$$

for $(t, y) \in T \times \mathcal{Y}$, where $\tilde{\lambda}_1(t, y, z) \leq \tilde{\lambda}_2(t, y, z) \leq \dots \leq \tilde{\lambda}_n(t, y, z)$ are the ordered roots of $\tilde{P}(t, y, z)$.

If $z \le 0$ and $r < 1$, then $\mathrm{Im}(z^r) = -|z|^r \sin r\pi \le 0$, by our demand. Therefore, $P(t,y)(x+z^r) \neq 0$, if $0 > \mathrm{Im}(x+z^r) = \mathrm{Im}(x) + \mathrm{Im}(z^r) = \mathrm{Im}(x) - |z|^r \sin r\pi$, since $P(t,y)$ is hyperbolic. Application of lemma 6.1.1 shows that

$$\tilde{P}(t,y,z)(x+z^r) \neq 0, \quad \text{when } \mathrm{Im}(x) < |z|^r \sin r\pi. \tag{6.16}$$

6.3.3. Next, for $1 \le j \le n$, let us expand by means of Taylor's formula

$$a_j(t+z\xi, y) = \sum_{|\alpha| \le k} \frac{\partial_t^\alpha a_j(t,y)}{\alpha!} z^{|\alpha|}\xi^\alpha + \tilde{a}_j(t,\xi,y,z),$$

where $z \in \mathbb{R}$, $\xi \in \mathbb{R}^m$, $t \in T$, $t + z\xi \in T$, and $y \in \mathcal{Y}$. Note that here the convexity of T is used. Then, we have, for a $0 \le \vartheta \le 1$,

$$\begin{aligned}
|\tilde{a}_j(t,\xi,y,z)| &= \left| \sum_{|\alpha|=k} \frac{\partial_t^\alpha a_j(t+\vartheta z\xi, y)}{\alpha!} z^{|\alpha|}\xi^\alpha - \sum_{|\alpha|=k} \frac{\partial_t^\alpha a_j(t,y)}{\alpha!} z^{|\alpha|}\xi^\alpha \right| \\
&\le \sum_{|\alpha|=k} |\partial_t^\alpha a_j(t+\vartheta z\xi, y) - \partial_t^\alpha a_j(t,y)| \frac{|z|^{|\alpha|}|\xi_1|^{\alpha_1}\cdots|\xi_m|^{\alpha_m}}{\alpha!} \\
&\le A|z|^{k+\delta}|\xi|^{k+\delta},
\end{aligned}$$

for a positive constant A, by the assumptions of the theorem and (6.12) and since $|\xi_1|^{\alpha_1}\cdots|\xi_m|^{\alpha_m} \le (\max_{1\le j\le m}|\xi_j|)^{|\alpha|} \le K|\xi|^{|\alpha|}$ (equivalence of norms). If, moreover, $|z| \le 1$ and $|\xi| \le 1$, this gives

$$|\tilde{a}_j(t,\xi,y,z)| \le A|z|^{nr}|\xi|^{nr}, \tag{6.17}$$

since $nr = \min\{n, k+\delta\}$.

6.3.4. Let U be an open relative-compact subset of T, and define

$$Q(t,\xi,y,z)(x) = \left(1 + z^r\frac{\partial}{\partial x}\right)^{n-1} \left(x^n + \sum_{j=1}^n \left(\sum_{|\alpha|\le k} \frac{\partial_t^\alpha a_j(t,y)}{\alpha!} z^{|\alpha|}\xi^\alpha \right) x^{n-j} \right).$$

Now, (6.14) yields that all roots of $\tilde{P}(t+z\xi,y,z)$ are distinct for $z > 0$ or $z \in \mathbb{R}\backslash\{0\}$ and $r = 1$, and the difference between two of these roots does not depend on ξ, because $t + z\xi$ which plays now the role of t in (6.14) does not appear in the right-hand side of (6.14). Observe that

$$\begin{aligned}
Q(t,\xi,y,z)(x) &= \left(1 + z^r\frac{\partial}{\partial x}\right)^{n-1} \left(x^n + \sum_{j=1}^n \left(a_j(t+z\xi,y) - \tilde{a}_j(t,\xi,y,z)\right) x^{n-j} \right) \\
&= \tilde{P}(t+z\xi,y,z)(x) - \left(1 + z^r\frac{\partial}{\partial x}\right)^{n-1} \sum_{j=1}^n \tilde{a}_j(t,\xi,y,z)x^{n-j}.
\end{aligned}$$

If we choose z and ξ sufficiently small, then, by (6.17), we can arrange $\tilde{a}_j(t,\xi,y,z)$ to be small enough such that all of the roots of $Q(t,\xi,y,z)$ are still distinct, see lemma 6.1.3. In other words, there are positive constants δ_0 and δ_1 such that $Q(t,\xi,y,z)(x) = 0$ has only simple roots, for (t,ξ,y) in

$$\Omega(U;\delta_1) := \{(t,\xi,y) \in U \times \mathbb{R}^m \times \mathcal{Y} : |\xi| \le \delta_1\},$$

if $0 < z \le \delta_0$ or $0 < |z| \le \delta_0$ and $r = 1$.

Furthermore, we know that all roots of $\tilde{P}(t+z\xi,y,z)$ are not only simple but also real. And the space of hyperbolic polynomials of degree n having only simple roots is open in the space of hyperbolic polynomials of degree n, see theorem 3.1.2. Consequently, if we put $z = \delta_0$ (or $|z| = \delta_0$ in the case $r = 1$), we can modify δ_1 such that all roots of $Q(t,\xi,y,\delta_0)$ (or $Q(t,\xi,y,\pm\delta_0)$ for $r = 1$) are real, for $(t, \xi, y) \in \Omega(U; \delta_1)$. But this implies that $Q(t, \xi, y, z)$ is hyperbolic, whenever $(t, \xi, y) \in \Omega(U; \delta_1)$ and $0 < z \leq \delta_0$ or $0 < |z| \leq \delta_0$ for $r = 1$, since all its roots are simple under these conditions and it is depending continuously on z (recall the description of the space of hyperbolic polynomials of degree n in section 3.1). Note that $Q(t, \xi, y, 0) = P(t, y)$ whose roots are all real by assumption. Summarizing we find that there are positive constants δ_0 and δ_1 such that $Q(t, \xi, y, z)(x) = 0$ has only real roots, for $(t, \xi, y) \in \Omega(U; \delta_1)$, if $0 \leq z \leq \delta_0$ or $-\delta_0 \leq z \leq \delta_0$ and $r = 1$.

Therefore, we can write

$$Q(t, \xi, y, z)(x + z^r) = \prod_{j=1}^{n} (x - \Lambda_j(t, \xi, y, z)),$$

for $(t, \xi, y) \in \Omega(U; \delta_1)$ and $0 \leq z \leq \delta_0$, where we assume that $\Lambda_1(t, \xi, y, z) \leq \Lambda_2(t, \xi, y, z) \leq \cdots \leq \Lambda_n(t, \xi, y, z)$.

6.3.5. Consider the second term on the right-hand side of

$$\begin{aligned} Q(t, \xi, y, z)(x + z^r) &= \tilde{P}(t + z\xi, y, z)(x + z^r) \\ &\quad - \left(1 + z^r \frac{\partial}{\partial x}\right)^{n-1} \sum_{j=1}^{n} \tilde{a}_j(t, \xi, y, z)(x + z^r)^{n-j}, \end{aligned} \tag{6.18}$$

for (t, ξ, y) and z as just before. By expanding and ordering the expression, it turns out to be a polynomial in x, where the coefficient of x^i, which we want to denote by b_{n-i} in view of lemma 6.1.3, has the following form

$$b_{n-i} = \sum_{j=1}^{n-i} \sum_{k=i}^{n-j} \frac{k!}{i!} \binom{n-j}{k} \binom{n-1}{k-i} z^{(n-i-j)r} \tilde{a}_j(t, \xi, y, z).$$

Using (6.17), we find

$$|b_{n-i}| \leq \sum_{j=1}^{n-i} \sum_{k=i}^{n-j} \frac{k!}{i!} \binom{n-j}{k} \binom{n-1}{k-i} z^{(n-i-j)r} A z^{nr} |\xi|^{nr}.$$

Hence, $|b_{n-i}|^{\frac{1}{n-i}}$ and $|b_{n-i}|^{\frac{1}{n}}$ are bounded from above by z^r multiplied by a constant factor, if $(t, \xi, y) \in \Omega(U; \delta_1)$ and $0 \leq z \leq \delta_0$. Therefore, the application of lemma 6.1.3 gives

$$|\Lambda_j(t, \xi, y, z) - (\tilde{\lambda}_j(t + z\xi, y, z) - z^r)| \leq k z^r,$$

with a constant k, since $\tilde{\lambda}_j(t + z\xi, y, z) - z^r$ (taking the part of α_j^0 in lemma 6.1.3) is continuous on the relative-compact set $\Omega(U; \delta_1) \times [0, \delta_0]$ and thus bounded on it. Summarizing we have found that there is a constant $c > 0$ such that

$$|\Lambda_j(t, \xi, y, z) - \tilde{\lambda}_j(t + z\xi, y, z)| \leq c z^r, \quad 1 \leq j \leq n, \tag{6.19}$$

for $(t, \xi, y) \in \Omega(U; \delta_1)$ and $0 \leq z \leq \delta_0$.

6.3.6. We claim that

$$Q(t, \xi, y, z)(x + z^r) \neq 0, \tag{6.20}$$

for $(t, \xi, y) \in \Omega(U; \delta_1)$, $\mathrm{Im}(x) < 0$, and $-\delta_0 \leq z \leq \delta_0$, modifying δ_0 and δ_1, if necessary. In fact, if $r = 1$ or $0 \leq z \leq \delta_0$, then the equation $Q(t, \xi, y, z)(x + z^r) = 0$ has only real roots, as we have found in 6.3.4. Still to investigate is the case, when $r < 1$ and $-\delta_0 \leq z < 0$. Then, we see, by (6.16), that $\tilde{P}(t + z\xi, y, z)(x + z^r) \neq 0$, for $\mathrm{Im}(x) < 0$, since $0 < |z|^r \sin r\pi$. Suppose that $Q(t, \xi, y, z)(x + z^r) = 0$ had a root Λ with $\mathrm{Im}(\Lambda) < 0$. In view of (6.18), by lemma 6.1.3, and by (6.17), we could find a root $\tilde{\lambda}$ of $\tilde{P}(t + z\xi, y, z)(x + z^r) = 0$ such that $|\Lambda - \tilde{\lambda}| = o(|z|^{c_1} |\xi|^{c_2})$ with positive constants c_1 and c_2. By shrinking δ_0 and δ_1, we could arrange $\tilde{\lambda}$ to lie in $\{x \in \mathbb{C} : \mathrm{Im}(x) < 0\}$, a contradiction. This proves the claim.

6.3.7. For $x \in \mathbb{R}$, $0 < z \leq \frac{\delta_0}{2}$, and $(t, \xi, y) \in \Omega(U; \delta_1)$, we can localize in the following sense:

$$Q(t, \xi, y, z + s\zeta)(x + z^{r-1}s\tau + (z + s\zeta)^r) = s^\mu \left(Q_{(x,z;t,\xi,y)}(\tau, \zeta) + o(1) \right),$$

as $s \searrow 0$, where $Q_{(x,z;t,\xi,y)}(\tau, \zeta) \not\equiv 0$ in (τ, ζ). The polynomial $Q_{(x,z;t,\xi,y)}(\tau, \zeta)$ is homogeneous in (τ, ζ) of degree μ ($\mu = 0$ is allowed) and satisfies

$$Q_{(x,z;t,\xi,y)}(-1, 0) \neq 0 \quad \text{and} \quad Q_{(x,z;t,\xi,y)}(\tau, \zeta) \neq 0, \text{ if } \mathrm{Im}(\tau) < 0 \text{ and } \zeta \in \mathbb{R}. \tag{6.21}$$

In fact, $Q(t, \xi, y, \tilde{z})(z^{r-1}\tilde{x} + \tilde{z}^r)$ is real analytic in $(\tilde{x}, \tilde{z})$ and microhyperbolic with respect to $(-1, 0) \in \mathbb{R}^2$ near the fixed value $(\tilde{x}, \tilde{z}) = (z^{1-r}x, z)$. Microhyperbolicity is seen as follows: By (6.20), we find

$$\begin{aligned} &Q(t, \xi, y, z + s\zeta)(x + z^{r-1}s\tau + (z + s\zeta)^r) \\ &\qquad = Q(t, \xi, y, z + s\zeta)(x + z^{r-1}s\mathrm{Re}(\tau) + iz^{r-1}s\mathrm{Im}(\tau) + (z + s\zeta)^r) \neq 0, \end{aligned}$$

since $\mathrm{Im}(x + z^{r-1}s\tau) = z^{r-1}s\mathrm{Im}(\tau) < 0$. (The part of the parameter t in (6.10) is played here by $-z^{r-1}s\mathrm{Im}(\tau)$). Lemma 6.2.2 yields the existence of the localization and (6.21). Note that $Q_{(x,z;t,\xi,y)}(\tau, \zeta)$ can be defined and fulfills (6.21), if $r = 1$ and $z = 0$, too.

6.3.8. We define

$$f(s, \zeta, (x, t, \tau, \xi, y, z)) = Q(t, \xi, y, z + s\zeta)(x + z^{r-1}s\tau + (z + s\zeta)^r),$$

for $s \in \mathbb{C}$ with $\mathrm{Im}(s) \leq 0$ and $|s| \leq s_0$, $\tau \in [\frac{1}{2}, \infty)$, $\zeta \in [0, 1]$, $x \in \mathbb{C}$ with $\mathrm{Im}(x) \leq 0$, $(t, \xi, y) \in \Omega(U; \delta_1)$, and $z \in (0, \epsilon]$, where $0 < s_0 \leq \frac{\delta_0}{2}$ and $0 < \epsilon \leq \frac{\delta_0}{2}$. This function is clearly continuous wherever it is defined. In the following consideration let us treat separately the cases, where $r < 1$ and $r = 1$:

6.3.9. Let $r < 1$. We are going to check now, whether this f satisfies the assumptions of lemma 6.1.4. The function f is clearly holomorphic in s, for $\mathrm{Im}(s) < 0$ (it corresponds to (i) in lemma 6.1.4). We find $f(s, \zeta, (x, t, \tau, \xi, y, z)) \neq 0$, if $\mathrm{Im}(x) < 0$ and $s \in \mathbb{R}$, by (6.20), since then $z + s\zeta \in [-\delta_0, \delta_0]$ and $\mathrm{Im}(x + z^{r-1}s\tau) = \mathrm{Im}(x) < 0$. This corresponds to (ii), since $\{x \in \mathbb{C} : \mathrm{Im}(x) < 0\}$ is dense in $\{x \in \mathbb{C} : \mathrm{Im}(x) \leq 0\}$. With respect to condition (iii) let us assert the following: For all $K > 0$ there is an $\epsilon > 0$ such that $f(s, \zeta, (x, t, \tau, \xi, y, z)) \neq 0$, if $|s| = s_0$, $|x| \leq K$, and $z \in (0, \epsilon]$. For, choosing ϵ small, makes z^{r-1} large which in turn makes $f(s, \zeta, (x, t, \tau, \xi, y, z))$ large, since it is a monic polynomial in $x + z^{r-1}s\tau + (z + s\zeta)^r$, and s is determined by $|s| = s_0$. To condition (iv) corresponds: $f(s, 0, (x, t, \tau, \xi, y, z)) \neq 0$ for $\mathrm{Im}(s) < 0$. That follows from (6.20), since $\mathrm{Im}(x + z^{r-1}s\tau) = \mathrm{Im}(x) + z^{r-1}\tau\mathrm{Im}(s) < 0$. The parts of M and U in lemma 6.1.4 are played here by $[0, 1]$ and $\{x \in \mathbb{C} : \mathrm{Im}(x) \leq 0, |x| \leq K\} \times \Omega(U; \delta_1) \times [\frac{1}{2}, \infty) \times (0, \epsilon]$, respectively. Thus, all assumptions are fulfilled, and we get

$$Q(t, \xi, y, z + s\zeta)(x + z^{r-1}s\tau + (z + s\zeta)^r) \neq 0, \tag{6.22}$$

if $r < 1$, $\mathrm{Im}(s) < 0$ and $|s| \leq s_0$, $\tau \in [\frac{1}{2}, \infty)$, $\zeta \in [0, 1]$, $\mathrm{Im}(x) \leq 0$ and $|x| \leq K$, $(t, \xi, y) \in \Omega(U; \delta_1)$, and $z \in (0, \epsilon]$.

6.3.10. The case, where $r = 1$, is slightly more difficult. Let $\tau_0 \geq 1$ be fixed. We claim that, for all $(x_0, t_0, \xi_0, y_0) \in \mathbb{R} \times U \times \mathbb{R}^m \times \mathcal{Y}$ with $|\xi_0| \leq \frac{\delta_1}{2}$, there is a $c' > 0$ such that

$$Q_{(x_0,0;t_0,\xi_0,y_0)}(\tau_0, \zeta) \neq 0,$$

if $\zeta \in [0, c']$. This is seen as follows: We have $Q_{(x_0,0;t_0,\xi_0,y_0)}(\tau_0, 0) = (-\tau_0)^\mu Q_{(x_0,0;t_0,\xi_0,y_0)}(-1, 0) \neq 0$, by (6.21), and continuity in the second variable gives the statement. The constant $c' = c'(\tau_0)$ depends on τ_0. But, if $\zeta \in (0, c'(1)]$, where $c'(1)$ denotes the constant c' associated to $\tau_0 = 1$, then $Q_{(x_0,0;t_0,\xi_0,y_0)}(\tau_0, \zeta) = \tau_0^\mu Q_{(x_0,0;t_0,\xi_0,y_0)}(1, \tau_0^{-1}\zeta) \neq 0$, since $\tau_0^{-1}\zeta \leq \zeta \leq c'(1)$. That means that we can choose $c' = c'(1)$, for all $\tau_0 \geq 1$, with the effect that c' is no longer depending on τ_0.

We assert that there exist $s_0 > 0$, $\epsilon > 0$, and a neighborhood V of y_0 in $\mathcal{Y}$ such that

$$Q(t, \xi, y, z + s\zeta)(x + s\tau + (z + s\zeta)) \neq 0, \tag{6.23}$$

if $\mathrm{Im}(s) \leq 0$ and $|s| = s_0$, $\tau \in [\tau_0 - \epsilon, \tau_0 + \epsilon]$, $\zeta \in [0, c']$, $|x - x_0| < \epsilon$, $(t, \xi, y) \in T \times \mathbb{R}^m \times V$ with $|t - t_0| < \epsilon$ and $|\xi - \xi_0| < \epsilon$, and $z \in [0, \epsilon]$. For we can write:

$$Q(t, \xi, y, z + s\zeta)(x + s\tau + (z + s\zeta)) = \sum_{j=0}^{\mu_0} Q_j(t, \xi, y, z, \tau, \zeta)(x) s^j + o(s^{\mu_0}),$$

as $s \to 0$, where $Q_j(t_0, \xi_0, y_0, 0, \tau, \zeta)(x_0)$ equals $Q_{(x_0,0;t_0,\xi_0,y_0)}(\tau, \zeta)$ for $j = \mu_0$ and vanishes identically for all $j < \mu_0$. Then assertion (6.23) follows by continuity.

Let us now apply lemma 6.1.4 again: Here (6.23) corresponds to condition (iii); (i), (ii), and (iv) are obvious, since (6.20) holds for $r = 1$, too. Therefore,

$$Q(t, \xi, y, z + s\zeta)(x + z^{r-1}s\tau + (z + s\zeta)^r) \neq 0, \tag{6.24}$$

if $r = 1$, $\mathrm{Im}(s) < 0$ and $|s| \leq s_0$, $\tau \in [\tau_0 - \epsilon, \tau_0 + \epsilon]$, $\zeta \in [0, c']$, $\mathrm{Im}(x) \leq 0$ and $|x - x_0| < \epsilon$, $(t, \xi, y) \in T \times \mathbb{R}^m \times V$ with $|t - t_0| < \epsilon$ and $|\xi - \xi_0| < \epsilon$, and $z \in [0, \epsilon]$.

6.3.11. Finally, we put together what we have found separately in the cases $r < 1$ and $r = 1$. Suppose that $K > 0$ and $\tau_0 \geq 1$ are given. By the considerations above, we find constants c', s_0, ϵ, and δ_1 and neighborhoods V of y_0 for all $(x_0, t_0, \xi_0, y_0) \in \{x \in \mathbb{R} : |x| \leq K\} \times U \times \{\xi \in \mathbb{R}^m : |\xi| \leq \frac{\delta_1}{2}\} \times \mathcal{Y}$ such that (6.24) holds. Since $\overline{U}$ and $\mathcal{Y}$ are compact, we can get rid of their dependence on (x_0, t_0, ξ_0, y_0) and state, consequently: For all $K > 0$ there are positive constants c', s_0, ϵ, and δ_1 such that

$$Q(t, \xi, y, z + s\zeta)(x + z^{r-1}s\tau + (z + s\zeta)^r) \neq 0, \tag{6.25}$$

if $r \leq 1$, $\mathrm{Im}(s) < 0$ and $|s| \leq s_0$, $\tau \in [\tau_0 - \epsilon, \tau_0 + \epsilon]$ with $\tau_0 \geq 1$, $\zeta \in [0, c']$, $\mathrm{Im}(x) \leq 0$ and $|x| \leq K$, $(t, \xi, y) \in \Omega(U; \delta_1)$, and $z \in (0, \epsilon]$.

We claim that this implies that, for all $\tau_0 \geq 1$,

$$Q_{(x,z;t,\xi,y)}(\tau_0, \zeta) \neq 0, \tag{6.26}$$

if $x \in \mathbb{R}$ and $|x| < K$, $z \in (0, \epsilon)$, $(t, \xi, y) \in \Omega(U; \delta_1)$, and $\zeta \in [0, c']$. In fact, assume that there exist $x_0 \in \mathbb{R}$ with $|x_0| < K$, $z_0 \in (0, \epsilon)$, $(t_0, \xi_0, y_0) \in \Omega(U; \delta_1)$, and $\zeta_0 \in [0, c']$ such that $Q_{(x_0,z_0;t_0,\xi_0,y_0)}(\tau_0, \zeta_0) = 0$. On the other hand consider

$$Q_{(x,z_0;t_0,\xi_0,y_0)}(\tau_0, \zeta_0) = s^{-\mu} Q(t_0, \xi_0, y_0, z_0 + s\zeta_0)(x + z_0^{r-1}s\tau_0 + (z_0 + s\zeta_0)^r) - o(1),$$

as $s \searrow 0$. By (6.25), we obtain

$$|s^{-\mu} Q(t_0, \xi_0, y_0, z_0 + s\zeta_0)(x + z_0^{r-1}s\tau_0 + (z_0 + s\zeta_0)^r)| > |o(1)|,$$

for sufficiently small s. Application of Rouché's theorem 2.2.2 yields that there are no roots of $Q_{(x,z_0;t_0,\xi_0,y_0)}(\tau_0, \zeta_0) = 0$ on the boundary of $\{x \in \mathbb{C} : \mathrm{Im}(x) \leq 0, |x| \leq K\}$, in contradiction to our assumption.

6.3.12. This enables us finally to prove the theorem. Since $0 < z < \epsilon \leq \frac{\delta_0}{2}$ and $z + s\zeta \leq \delta_0$ for small s, we can write

$$\begin{aligned} 0 &= Q(t,\xi,y,z+s\zeta)(\Lambda_j(t,\xi,y,z+s\zeta) + (z+s\zeta)^r) \\ &= s^\mu \left(Q_{(\Lambda_j(t,\xi,y,z),z;t,\xi,y)}(z^{1-r}s^{-1}(\Lambda_j(t,\xi,y,z+s\zeta) - \Lambda_j(t,\xi,y,z)),\zeta) + o(1)\right), \end{aligned} \tag{6.27}$$

as $s \searrow 0$, if $(t,\xi,y) \in \Omega(U;\delta_1)$, $\zeta \in [0,c']$, and $z \in (0,\epsilon)$. Note that μ is depending on (t,ξ,y,z) and on j. If in (6.27) the parameter s approaches 0, then we find

$$Q_{(\Lambda_j(t,\xi,y,z),z;t,\xi,y)}(\underbrace{z^{1-r}s^{-1}(\Lambda_j(t,\xi,y,z+s\zeta) - \Lambda_j(t,\xi,y,z))}_{\to z^{1-r}\frac{\partial}{\partial s}\big|_{s=0}\Lambda_j(t,\xi,y,z+s\zeta)},\zeta) \to 0.$$

But by (6.26) this is impossible, if $z^{1-r}\left.\frac{\partial}{\partial s}\right|_{s=0}\Lambda_j(t,\xi,y,z+s\zeta) \geq 1$. So we have found

$$\left.\frac{\partial}{\partial s}\right|_{s=0}\Lambda_j(t,\xi,y,z+s\zeta) < z^{r-1}, \tag{6.28}$$

for $(t,\xi,y) \in \Omega(U;\delta_1)$, $\zeta \in [0,c']$, and $z \in (0,\epsilon)$.

Let us collect the estimates we have found: Note that $\Lambda_j(t,\xi,y,0) = \lambda_j(t,y)$ and use (6.15), (6.19), and (6.28),

$$\lambda_j(t+z\xi,y) - \lambda_j(t,y) \leq K'z^r,$$

for $(t,\xi,y) \in \Omega(U;\delta_1)$ and $z \in [0,\epsilon]$, where K' is a positive constant. Exchanging $t + z\xi$ and t in the previous calculation, and recalling the compactness of $\overline{U}$ and $\mathcal{Y}$, we obtain that there is a positive constant C' such that

$$|\lambda_j(t_1,y) - \lambda_j(t_2,y)| \leq C'|t_1 - t_2|^r,$$

for $t_1, t_2 \in U$ and $y \in \mathcal{Y}$. This establishes (6.13) and completes the proof of the theorem.

6.3.13. Remark. Note that theorem 6.3.1 implies Bronshtein's result in theorem 5.5.13: Let T be an open interval in $\mathbb{R}$ and suppose that the partial derivatives $\frac{\partial^i}{\partial t^i}a_j(t,y)$ $(0 \leq i \leq n; 1 \leq j \leq n)$ are continuous on $T \times \mathcal{Y}$. It follows that $\frac{\partial^{n-1}}{\partial t^{n-1}}a_j(t,y)$ satisfies a Lipschitz condition with positive C and $\delta = 1$, for all $1 \leq j \leq n$. So, by theorem 6.3.1, for each open relative-compact $U \subseteq T \times \mathcal{Y}$ there is a constant C_U such that the roots $\lambda_1(t,y) \leq \cdots \leq \lambda_n(t,y)$ fulfill

$$|\lambda_j(t,y) - \lambda_j(t',y)| \leq C_U|t - t'|$$

for all $t, t' \in U$ and $y \in \mathcal{Y}$. But this estimate implies that indeed $\frac{\partial}{\partial t}\lambda_j(t,y)$ is bounded on U.

6.3.14. Remark. In [**KP08**] a special case of theorem 6.3.1 is proved: Let $P(t)(x) = x^n + \sum_{i=1}^n (-1)^i a_i(t) x^{n-i}$ be a hyperbolic polynomial, where a_i are real analytic functions in an open subset $U \subseteq \mathbb{R}^m$. Then the increasingly ordered roots $\lambda_1(t) \leq \cdots \leq \lambda_n(t)$ are locally Lipschitz. The method of proof differs significantly from the argumentation in [**Bro79**] and [**Wak86**], respectively. Basically, the proof is obtained by a reduction to the two parameter case and by a careful study of a desingularization of singularities of the zeros of P.

Another theorem in [**KP08**] states that after suitable blowing-ups of the space of parameters one can write locally the roots of P as analytic functions of parameters.

CHAPTER 7

An application of Bronshtein's result

7.1. Twice differentiable parameterization of the roots

The result of Bronshtein and Wakabayashi on the boundedness of the derivatives of the roots of a curve of polynomials with coefficients in C^n, where n is the polynomial degree, can be used to construct a twice differentiable parameterization of the roots of any curve of hyperbolic polynomials with coefficients in C^{3n}. This conclusion is best possible, since the second derivatives may be unbounded even if the coefficients are smooth; compare with 4.1.4 and 4.1.5.

7.1.1. The following theorem is due to Kriegl, Losik, and Michor [**KLM04**]. However, the first statement in theorem 7.1.1 has been proved by Mandai [**Man85**] before. Mandai's proof is essentially the same as the proof presented in [**KLM04**]. It relies on the boundedness result of Bronshtein, theorem 5.5.13.

Theorem. *Consider a continuous curve of hyperbolic polynomials*

$$P(t)(x) = x^n - a_1(t)x^{n-1} + \cdots + (-1)^n a_n(t) \qquad (t \in \mathbb{R}).$$

Then there is a continuous parameterization $x = (x_1, \dots, x_n) : \mathbb{R} \to \mathbb{R}^n$ of the roots of P. Moreover,

(1) *If all a_i are of class C^{2n}, then any differentiable parameterization of the roots $x : \mathbb{R} \to \mathbb{R}^n$ is actually C^1.*
(2) *If all a_i are of class C^{3n}, then the parameterization of the roots $x : \mathbb{R} \to \mathbb{R}^n$ may be chosen twice differentiable.*

Proof. We know that the parameterization of the roots by ordering them increasingly, $x_1(t) \le \cdots \le x_n(t)$, is continuous (theorem 2.3.2). By replacing x by $y = x - \frac{a_1(t)}{n}$, we may assume that $a_1 = 0$.

According to the strong multiplicity lemma 4.2.8, for $m \ge n$ the following statements are equivalent:

(i) $a_k(t) = t^k a_{k,k}(t)$ for a C^{m-k} function $a_{k,k}$, for all $2 \le k \le n$;
(ii) $a_2(t) = t^2 a_{2,2}(t)$ for a C^{m-2} function $a_{2,2}$.

To the proof of (1): Let all a_i be C^{2n}. We choose a fixed t, say $t = 0$. We shall repeat with slight modifications the proof of theorem 4.3.2.

If $a_2(0) = 0$, then $a_2(t) = t^2 a_{2,2}(t)$, and so by the variant of the multiplicity lemma described above we have $a_k(t) = t^k a_{k,k}(t)$ for C^n functions $a_{k,k}$, for all $2 \le k \le n$. Consider the following C^n curve of hyperbolic polynomials

$$P^1(t)(z) = z^n + a_{2,2}(t)z^{n-2} - a_{3,3}(t)z^{n-3} + \cdots + (-1)^n a_{n,n}(t) \qquad (7.1)$$

which satisfies $P^1(t)(z) = t^{-n}P(t)(tz)$. Hence $z \mapsto tz$ gives for $t \ne 0$ a bijective correspondence between the roots z of P^1 and the roots x of P with correct multiplicities. Moreover, parameterizations z which are continuous at $t = 0$ correspond to parameterizations x which are differentiable at $t = 0$. By theorem 5.5.13 we may choose the parameterization $z = (z_1, \dots, z_n)$ differentiable with locally bounded derivative. Then the corresponding parameterization $t \mapsto x(t) := tz(t)$ is differentiable with derivative $x'(t) = tz'(t) + z(t)$ which is continuous at $t = 0$ with $x'(0) = z(0)$.

If $a_2(0) \neq 0$, then we apply the splitting lemma 4.2.3: We can factorize $P(t) = P_1(t) \cdots P_k(t)$ for t in a neighborhood of 0 and some integer $k > 1$, where the P_i have again C^{2n} coefficients and where each $P_i(0)$ has all roots equal to, say, c_i, and where the c_i are distinct. By the argument above applied to each P_i separately, there is a differentiable parameterization $x = (x_1, \dots, x_n)$ of roots whose derivative x' is continuous at $t = 0$. Furthermore, if $x_j(0)$ is a root of $P_i(0)$, then $x_j'(0)$ is a root of the polynomial $P_i^1(0)$ which depends only on P_i. We shall use this for arbitrary t below.

Now we shall prove that any differentiable parameterization $y = (y_1, \dots, y_n)$ of roots of P has continuous derivative y' at $t = 0$. Let $1 \leq i \leq n$ be fixed. For $t_m \to 0$ there are $1 \leq k_m \leq n$ such that $y_i(t_m) = x_{k_m}(t_m)$. Choose a subsequence of $(t_m)_m$, again denoted by $(t_m)_m$, such that $y_i(t_m) = x_k(t_m)$ for some fixed k and all m. Then, by the argument at the end of the last paragraph, we also have $y_i'(t_m) = x_{j_m}'(t_m)$ for some $1 \leq j_m \leq n$ with $x_{j_m}(t_m) = x_k(t_m) = y_i(t_m)$. Passing again to a subsequence, we find a fixed j such that $y_i(t_m) = x_j(t_m)$ and $y_i'(t_m) = x_j'(t_m)$. Consequently,

$$y_i(0) = \lim_{m\to\infty} y_i(t_m) = \lim_{m\to\infty} x_j(t_m) = x_j(0),$$

and

$$y_i'(0) = \lim_{m\to\infty} \frac{y_i(t_m) - y_i(0)}{t_m} = \lim_{m\to\infty} \frac{x_j(t_m) - x_j(0)}{t_m} = x_j'(0),$$

and so $y_i'(t_m) = x_j'(t_m) \to x_j'(0) = y_i'(0)$. Since $t = 0$ was arbitrary, this shows that any differentiable parameterization of the roots of P, which exists by theorem 4.3.2, is indeed C^1, and (1) is proved.

To the proof of (2): Let all a_i be C^{3n}. Remember that $a_1 = 0$.

We start with a preliminary consideration. Choose a fixed t, say $t = 0$. If $a_2(0) = 0$, then we consider again the hyperbolic polynomials $P^1(t)$ in (7.1) which now form a C^{2n} curve. By (1) its roots can be parameterized by a C^1 curve $t \mapsto z(t) = (z_1(t), \dots, z_n(t))$. Then, $x(t) := tz(t)$ parameterizes the roots of $P(t)$ now with continuous derivative $x'(t) = tz'(t) + z(t)$ which is differentiable at $t = 0$ with

$$x''(0) = \lim_{t\to 0} \frac{tz'(t) + z(t) - z(0)}{t} = \lim_{t\to 0} z'(t) + \lim_{t\to 0} \frac{z(t) - z(0)}{t} = 2z'(0).$$

We show by induction on the polynomial degree n that for fixed intervals $I \subseteq \mathbb{R}$ there exists a twice differentiable parameterization y of the roots of P on I.

For $n = 1$ the only root equals the single coefficient. So let us assume the assertion is true for degrees strictly smaller than n.

Let $t_0 \in I$ be such that $a_2(t_0) \neq 0$. By the splitting lemma 4.2.3 we may factorize $P(t) = P_1(t) \cdots P_k(t)$ for some integer $k > 1$ and all t in a neighborhood $I_1 \subseteq I$ of t_0, where the P_i have again C^{3n} coefficients and where each $P_i(t_0)$ has all roots equal to, say, c_i, and where the c_i are distinct. The P_i have smaller degrees than P, so by induction hypothesis there is on I_1 a twice differentiable parameterization of the roots of each P_i.

Let now $a_2(t) \neq 0$ for all $t \in I$. We have seen that then for all $t \in I$ there exist twice differentiable parameterizations of the roots defined on open subintervals of I. Obviously we may apply Zorn's lemma to obtain a twice differentiable parameterization y on some maximal open subinterval $I_1 \subseteq I$. Suppose for contradiction that $I_1 \subsetneq I$ and let the, say, right endpoint t_0 of I_1 belong to I. Since $a_2(t_0) \neq 0$, there is a twice differentiable parameterization x of the roots in a neighborhood

$I_2 \subseteq I$ of t_0. Consider a sequence $(t_m)_{m\in\mathbb{N}} \subseteq I_1 \cap I_2$ with $t_m \nearrow t_0$. For every $m \in \mathbb{N}$ there exists a permutation π of $\{1,\dots,n\}$ such that $y_{\pi(i)}(t_m) = x_i(t_m)$ for all i. By passing to a subsequence, again denoted by $(t_m)_m$, we may assume that the permutation π does not depend on m. By passing again to a subsequence we can also assume that $y'_{\pi(i)}(t_m) = x'_i(t_m)$ and then again for a subsequence that $y''_{\pi(i)}(t_m) = x''_i(t_m)$ for all i and all m. So we are able to paste the parameterization $(y_{\pi(i)}(t))_i$ for $t < t_0$ with the parameterization $x(t)$ for $t \ge t_0$ to obtain a twice differentiable parameterization on an interval larger than I_1, a contradiction.

Now we consider the closed set

$$E = \{t \in I : a_2(t) = 0\} = \{t \in I : x_1(t) = \cdots = x_n(t) = 0\},$$

where $x_1,\dots,x_n$ denote the roots of P. Then $I\backslash E$ is open, thus a disjoint union of open intervals on which we have a twice differentiable parameterization x of the roots by the previous paragraph.

Consider next the set E' of all accumulation points of E. Then $I\backslash E' = (I\backslash E) \cup (E\backslash E')$ is again open and hence a disjoint union of open intervals. For each point $t_0 \in E\backslash E'$, i.e., isolated point of E, we have a twice differentiable local parameterization of roots $y_i(t)$ for $t \neq t_0$ (left and right of t_0), and there is a C^1 parameterization $x_k(t)$ for t near t_0 which is twice differentiable at t_0, by the preliminary consideration. Clearly, $y_i(t) \to x_1(t_0) = \cdots = x_n(t_0) = 0$ for $t \to t_0$ and for all i.

For a sequence $(t_m)_{m\in\mathbb{N}}$ with $t_m \searrow t_0$, by passing to a subsequence denoted equally, we may assume that $y'_i(t_m) = x'_{\pi(i)}(t_m) \to x'_{\pi(i)}(t_0)$ for a permutation π of $\{1,\dots,n\}$ not depending on m. Consequently, $y'_i(t)$ has at most $x'_1(t_0),\dots,x'_n(t_0)$ as cluster points for $t \searrow t_0$. Since y'_i satisfies the intermediate value theorem, $y'_i(t)$ converges for $t \searrow t_0$ with limit $x'_{\pi(i)}(t_0)$, since it does so along a sequence $(t_m)_m$ as above. By renumbering the y_i to the right of t_0 we may assume that $i = \pi(i)$. These arguments work similarly for the left side of t_0. We conclude that $y'_i(t) \to x'_i(t_0)$ for $t \to t_0$, so the parameterization y_i is C^1 near t_0 and still twice differentiable off t_0.

In order to get twice differentiability at t_0 also, we consider again the situation at the beginning of the last paragraph. Then we have

$$\frac{y'_i(t_m) - y'_i(t_0)}{t_m - t_0} = \frac{x'_{\pi(i)}(t_m) - x'_{\pi(i)}(t_0)}{t_m - t_0} \to x''_{\pi(i)}(t_0),$$

since the parameterization x_k is twice differentiable at t_0. Therefore, $\frac{y'_i(t)-y'_i(t_0)}{t-t_0}$ has at most $\{x''_j(t_0) : x'_j(t_0) = y'_i(t_0)\}$ as cluster points for $t \searrow t_0$. Since it satisfies the intermediate value theorem, it converges for $t \searrow t_0$ with limit $x''_{\pi(i)}(t_0)$, since it does so along a sequence $(t_m)_m$. We can argue similarly for the left-handed second derivative. Thus we may renumber those y_i for which the $y'_i(t_0)$ agree to the right of t_0 in such a way that the (one sided) second derivatives agree. Then the (twice) renumbered y_i are twice differentiable also at t_0.

That means we have constructed a twice differentiable parameterization of the roots of P on the open set $I\backslash E'$.

Now let $t' \in E'$, i.e., an accumulation point of E. Let F be the set of all $t \in I$ such that $x_1(t) = \cdots = x_n(t)$ and $x'_1(t) = \cdots = x'_n(t)$, where $x = (x_1,\dots,x_n)$ is a C^1 parameterization of the roots of P provided by (1). Then $t' \in F$, since each $x'_i(t')$ may be computed using only points in E. Let F' be the set of all accumulation points of F. Then we have $E' \subseteq F = (F\backslash F') \cup F' \subseteq E$.

Let first t' be an isolated point of F, i.e., $t' \in F\backslash F'$. Then again we have a local twice differentiable parameterization $t \mapsto y(t)$ of the roots for $t \neq t'$ (left and right of t'), since near t' there are only points of $I\backslash E'$. We still have a local C^1

parameterization x near t' which is twice differentiable at t', by the preliminary consideration. As above we can find a twice differentiable parameterization y of the roots of P on the open set $(I\backslash E') \cup (F\backslash F')$.

Finally we want two extend the parameterization $y = (y_1, \dots, y_n)$ obtained in the last paragraph to F'. Let t' be an accumulation point of F, i.e., $t' \in F'$. Again we are given a C^1 parameterization x near t' which is twice differentiable at t'. Then all $x_i(t')$ agree, all $x_i'(t')$ agree, and even all $x_i''(t')$ agree, since each $x_i''(t')$ may be computed using only points in F. Let us extend each y_i from $(I\backslash E') \cup (F\backslash F')$ by this single function on F' to the whole of $(I\backslash E') \cup (F\backslash F') \cup F' = (I\backslash E') = I$. We have to check that then each y_i is twice differentiable at t': For a sequence $(t_m)_{m\in\mathbb{N}}$ with $t_m \to t'$ we have, by passing to a subsequence,

$$y_i(t_m) = x_j(t_m) \to x_j(t') = x_i(t') = y_i(t'),$$

further

$$\frac{y_i(t_m) - y_i(t')}{t_m - t'} = \frac{x_j(t_m) - x_j(t')}{t_m - t'} \to x_j'(t') = x_i'(t'),$$

and finally

$$\frac{y_i'(t_m) - y_i'(t')}{t_m - t'} = \frac{x_j'(t_m) - x_j'(t')}{t_m - t'} \to x_j''(t') = x_i''(t').$$

This completes the induction. For $I = \mathbb{R}$ it yields the statement of (2). $\square$

Remark. Comparing this result with proposition 4.1.1, where we treated the quadratic case, shows that the differentiability assumptions for the curve of polynomials P in theorem 7.1.1 can possibly be improved.

Part 2

Lifting smooth curves over invariants

CHAPTER 8

The problem of lifting curves

8.1. A different point of view

8.1.1. In part 1 we considered monic polynomials

$$P(t)(x) = x^n - a_1(t)x^{n-1} + \cdots + (-1)^n a_n(t)$$

of fixed degree n having all roots real and being parameterized by t near 0 in $\mathbb{R}$ real analytically, smoothly or in a C^m way. And we investigated the problem of finding parameterizations $x_1(t), \ldots, x_n(t)$ of the roots of $P(t)$ with best possible differentiability properties. Note that in section 4.6 we treated additionally the cases when the coefficients and the roots of $P(t)$ are complex valued and when $P(t)$ is parameterized holomorphically by a complex parameter t. But let us restrict to the hyperbolic setting here.

8.1.2. Reformulation. The problem can be reformulated in the following way. Let the symmetric group S_n act in $\mathbb{R}^n$ by permuting the coordinates; they correspond to the roots of P. Consider the polynomial mapping $\sigma = (\sigma_1, \ldots, \sigma_n) : \mathbb{R}^n \to \mathbb{R}^n$ whose components are the elementary symmetric polynomials:

$$\sigma_i(x_1, \ldots, x_n) = \sum_{1 \le j_1 < \cdots < j_i \le n} x_{j_1} \cdots x_{j_i};$$

they correspond to the coefficients of P, by Vieta's formulas. Now the question is: Given a smooth curve $c : \mathbb{R} \to \sigma(\mathbb{R}^n) \subseteq \mathbb{R}^n$, is it possible to find a smooth *lift* $\bar{c} : \mathbb{R} \to \mathbb{R}^n$ of c, i.e., a smooth curve $\bar{c}$ satisfying $\sigma \circ \bar{c} = c$? The curve c corresponds to the curve P in the space of hyperbolic polynomials of degree n, which we may identify with $\sigma(\mathbb{R}^n)$, and the lift $\bar{c}$ corresponds to a parameterization of the roots of P.

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, two heads, "\sigma"] & \sigma(\mathbb{R}^n) \subseteq \mathbb{R}^n \\ & \mathbb{R} \arrow[u, "c"'] \arrow[ul, dotted, "\exists \bar{c}\,?"] \end{tikzcd}$$

8.1.3. Generalization. In this formulation the above problem suggests the following generalization. Consider an orthogonal representation of a compact Lie group G on a real finite dimensional Euclidean vector space V. Let $\sigma_1, \ldots, \sigma_n$ be a system of homogeneous generators for the algebra $\mathbb{R}[V]^G$ of invariant polynomials on V. Then the mapping $\sigma = (\sigma_1, \ldots, \sigma_n) : V \to \mathbb{R}^n$ induces an identification of the orbit space V/G with the semialgebraic set $\sigma(V) \subseteq \mathbb{R}^n$. A curve $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ in the orbit space V/G is called smooth, if it is smooth as a curve in $\mathbb{R}^n$. We shall see in remark 9.2.3(1) that the set $\sigma(V)$ does not depend on the choice of generators $\sigma_1, \ldots, \sigma_n$, hence this is well defined. Now we may ask: Given a smooth curve $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ in the orbit space, does there exist a smooth lift to V, i.e., a smooth curve $\bar{c} : \mathbb{R} \to V$ satisfying $\sigma \circ \bar{c} = c$?

$$\begin{tikzcd} V \arrow[r, two heads, "\sigma"] & \sigma(V) = V/G \subseteq \mathbb{R}^n \\ & \mathbb{R} \arrow[u, "c"'] \arrow[ul, dotted, "\exists \bar{c}\,?"] \end{tikzcd}$$

8.1.4. As in the case of choosing roots of polynomials, we will not just consider the smooth case, but instead we shall vary the differentiability conditions of the curve c during the treatment of this problem. In particular, we will investigate the lifting problem under real analyticity and finite differentiability.

8.2. The setting

8.2.1. Representations of compact Lie groups. Let G be a compact Lie group and let $\rho : G \to \mathrm{O}(V)$ be an orthogonal representation in a real finite dimensional Euclidean vector space V with inner product $\langle . \mid . \rangle$. By a classical theorem of Hilbert and Nagata the algebra $\mathbb{R}[V]^G$ of invariant polynomials on V is finitely generated, compare with lemma 9.3.3(1). So let $\sigma_1, \dots, \sigma_n$ be a system of homogeneous generators of $\mathbb{R}[V]^G$ with positive degrees $d_1, \dots, d_n$; assuming their homogeneity is no restriction. Let us consider the *orbit map*

$$\sigma = (\sigma_1, \dots, \sigma_n) : V \to \mathbb{R}^n.$$

Note that if $(y_1, \dots, y_n) = \sigma(v)$ for $v \in V$, then $(t^{d_1} y_1, \dots, t^{d_n} y_n) = \sigma(tv)$ for $t \in \mathbb{R}$. The image $\sigma(V)$ is a *semialgebraic* set, i.e., given by finitely many polynomial equations and inequalities, in the categorical quotient

$$V/\!\!/G := \{y \in \mathbb{R}^n : P(y) = 0 \text{ for all } P \in I\},$$

where I is the ideal of relations between $\sigma_1, \dots, \sigma_n$.

Under these conditions we have the following lemma.

8.2.2. Lemma. *In the above situation we have:*

(1) *σ is* proper*, i.e., pre-images of compact sets are compact.*
(2) *σ separates orbits of G.*
(3) *There is a map $\sigma' : V/G \to \mathbb{R}^n$ such that the following diagram commutes,*

$$\begin{tikzcd} V \arrow[r, "\sigma"] \arrow[d, two heads, "\pi"'] & \mathbb{R}^n \\ V/G \arrow[ur, dotted, "\sigma'"'] & \end{tikzcd}$$

and σ' is a homeomorphism onto its image.

Proof. To (1): Let $r(x) = \langle x \mid x \rangle$. Then $r \in \mathbb{R}[V]^G$. By the theorem of Hilbert and Nagata, see lemma 9.3.3, there is a polynomial $p \in \mathbb{R}[\mathbb{R}^n]$ such that $r(x) = p(\sigma(x))$. If $(x_m)_m \subseteq V$ is an unbounded sequence, then $(r(x_m))_m$ is unbounded. Therefore, $(p(\sigma(x_m)))_m$ is unbounded, and, since p is a polynomial, $(\sigma(x_m))_m$ is unbounded, too. With this insight we can conclude that any compact and hence bounded set in $\mathbb{R}^n$ must have a bounded pre-image under σ. By continuity of σ, it must be closed as well. Thus, σ is proper.

To (2): Let us choose two different orbits $G.x \neq G.y$ ($x, y \in V$); we have to show $\sigma(G.x) \neq \sigma(G.y)$. Consider the following map:

$$f : G.x \cup G.y \to \mathbb{R} \quad \text{with} \quad f(v) := \begin{cases} 0 & \text{for } v \in G.x \\ 1 & \text{for } v \in G.y \end{cases}.$$

This map is well defined, since if $G.x$ and $G.y$ have nonempty intersection then they agree completely. Both orbits are closed, so f is continuous. Furthermore, both

orbits and with them their union are compact, since G is compact. Therefore, by the Weierstrass approximation theorem, there exists a polynomial $p \in \mathbb{R}[V]$ such that
$$\|p - f\|_{G.x \cup G.y} = \sup\{|p(z) - f(z)| : z \in G.x \cup G.y\} < \frac{1}{10}.$$
Now we can average p over the group using the Haar measure dg on G to get a G-invariant function q on V:
$$q(v) := \int_G p(g.v)dg.$$
Note that since the action of G is linear, q is again a polynomial. Next let us check that q approximates f equally well. For $v \in G.x \cup G.y$, we have
$$\begin{aligned} |f(v) - q(v)| &= \left| \int_G f(g.v)dg - \int_G p(g.v)dg \right| \\ &\leq \int_G |f(g.v) - p(g.v)|dg \\ &\leq \frac{1}{10} \int_G dg = \frac{1}{10}. \end{aligned}$$
Recalling the definition of f we obtain
$$|q(v)| \leq \frac{1}{10} \qquad \text{for } v \in G.x$$
and
$$|1 - q(v)| \leq \frac{1}{10} \qquad \text{for } v \in G.y.$$
Therefore, $q(G.x) \neq q(G.y)$. Now $q \in \mathbb{R}[V]^G$ and can be expressed in the Hilbert generators $\sigma_1, \ldots, \sigma_n$. This implies that $\sigma(G.x) \neq \sigma(G.y)$.

To (3): The map $\sigma' : V/G \to \mathbb{R}^n : \pi(v) \mapsto \sigma(v)$ is well defined, since σ is G-invariant. By (2), σ' is injective and, with the quotient topology on V/G, continuous and proper. Then the statement follows from lemma 8.2.3. □

In the sequel we shall identify V/G and $\sigma(V)$ via the homeomorphism σ' given in lemma 8.2.2.

8.2.3. Lemma. *Suppose that X and Y are locally compact, Hausdorff spaces and that $f : X \to Y$ is bijective, continuous, and proper. Then f is a homeomorphism.*

Proof. (E.g. [**Bre93**]) By defining $\tilde{f}(\infty) = \infty$, f extends to a continuous map $\tilde{f} : X \cup \{\infty\} \to Y \cup \{\infty\}$ between the one point compactifications, since it is proper. For: Suppose $U \subseteq Y \cup \{\infty\}$ is open. If $U \subseteq Y$, then $\tilde{f}^{-1}(U) = f^{-1}(U)$ is open. In the other case $U = Y \cup \{\infty\} \backslash K$, where $K \subseteq Y$ is compact. Then $\tilde{f}^{-1}(U) = X \cup \{\infty\} \backslash f^{-1}(K)$ is open in $X \cup \{\infty\}$, since $f^{-1}(K)$ is compact and thus closed, by properness.

If $A \subseteq X$ is closed in X, then $A \cup \{\infty\}$ is closed in $X \cup \{\infty\}$ and hence compact. Then, $\tilde{f}(A \cup \{\infty\})$ is compact and hence closed in $Y \cup \{\infty\}$. Consequently, $f(A) = \tilde{f}(A \cup \{\infty\}) \cap Y$ is closed in Y. □

8.2.4. Functional structure on orbit spaces. Let a Lie group G act smoothly on a smooth manifold M. That is the smooth mapping $l : G \times M \to M : (g, x) \mapsto l(g, x) = l_g(x) = l^x(g) = g.x$ satisfies $l_g \circ l_h = l_{gh}$ and $l_e = \mathrm{id}_M$. Then the orbit space M/G is not generally again a smooth manifold. Yet, it still carries a functional structure induced by the smooth structure on M simply by calling a function $f : M/G \to \mathbb{R}$ smooth if and only if $f \circ \pi : M \to \mathbb{R}$ is smooth, where $\pi : M \to M/G$ is the quotient map. That means, the functional structure on M/G is determined completely by the smooth G-invariant functions on M. For compact Lie groups, the space of G-invariant C^∞-functions on V is characterized in the following theorem due to Gerald Schwarz:

8.2.5. THEOREM (Schwarz's theorem [**Sch75**]). *Consider a finite dimensional representation $\rho : G \to \mathrm{O}(V)$ of a compact Lie group G. Let $\sigma_1, \dots, \sigma_n$ be generators for the algebra $\mathbb{R}[V]^G$ of G-invariant polynomials on V. If $\sigma := (\sigma_1, \dots, \sigma_n) : V \to \mathbb{R}^n$, then*

$$\sigma^* : C^\infty(\mathbb{R}^n) \to C^\infty(V)^G$$

is surjective.

We shall give a proof of this theorem in section 9.1.

REMARK. In the early 1940's Whitney [**Whi43**] proved that every smooth even function $f : \mathbb{R} \to \mathbb{R}$ can be written as $f(x) = g(x^2)$, where g is smooth. In 1963 Glaeser [**Gla63a**] established the above theorem in the special case when ρ is the standard representation of the symmetric group S_n in $\mathbb{R}^n$, i.e., he showed that a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ which is invariant under permutation of the coordinates can be expressed as $f(x) = g(\sigma_1(x), \dots, \sigma_n(x))$, where g is smooth and the σ_i are the elementary symmetric functions.

In [**Mat77**] John N. Mather strengthens the conclusion of Schwarz's theorem. He proves that the mapping $\sigma^* : C^\infty(\mathbb{R}^n) \to C^\infty(V)^G$ is split surjective. A continuous linear mapping of topological vector spaces over $\mathbb{R}$ is said to be *split surjective* if there is a continuous $\mathbb{R}$-linear right inverse.

In [**Lun76**] D. Luna gives another generalization of Schwarz's theorem. He considers a reductive group acting linearly on $\mathbb{R}^k$, i.e., the action is completely reducible, see section 9.3. If $\sigma_1, \dots, \sigma_n$ is a generating set of the algebra of invariant polynomials and $C^\infty(\mathbb{R}^k; \sigma)$ denote the smooth functions on $\mathbb{R}^k$ which are constant on the fibers of $\sigma = (\sigma_1, \dots, \sigma_n)$, then $\sigma^* : C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^k; \sigma)$ is surjective. Mather's approach permits a generalization of Luna's result, too. In fact, he proves in [**Mat77**] that $\sigma^* : C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^k; \sigma)$ is split surjective.

Smooth invariants of a group action are a special example of the following much more general problem going back to Thom and Glaeser: Let M be a real analytic manifold and $\varphi : M \to \mathbb{R}^n$ a proper real analytic mapping. Suppose that $f : M \to \mathbb{R}$ is a C^∞ function. Under what conditions is f a composite $f = g \circ \varphi$, where g is a C^∞ function on $\mathbb{R}^n$? An obvious necessary condition is that f be constant on the fibers $\varphi^{-1}(x)$, where $x \in X := \varphi(M)$. A further necessary formal condition is that the Taylor expansions of f along any fiber $\varphi^{-1}(x)$ are the pullbacks of a formal power series centered at x. It is known that, if these two conditions are satisfied and $X = \varphi(M)$ is closed Nash subanalytic, then indeed f is a composite $f = g \circ \varphi$ with $g \in C^\infty(\mathbb{R}^n)$. Closed subanalytic sets are precisely the images of real analytic sets by proper real analytic mappings. Such a set X is called Nash if any point admits a neighborhood U such that $X \cap U$ is finite union of pure-dimensional subanalytic sets each of which lies in a closed analytic subset of U of the same dimension. The class of Nash subanalytic sets includes all closed semianalytic sets. This problem has been studied by Bierstone and Milman, see [**BM00**] and references therein.

8.2.6. DEFINITION. Let $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a smooth curve in the orbit space; smooth as curve in $\mathbb{R}^n$. A curve $\bar{c} : \mathbb{R} \to V$ is called *lift* of c to V, if $\sigma \circ \bar{c} = c$ holds.

8.2.7. Independence of the choice of generators. *The problem of lifting smooth curves over invariants is independent of the choice of a system of homogeneous generators of $\mathbb{R}[V]^G$ in the following sense:* Suppose $\sigma_1, \dots, \sigma_n$ and $\tau_1, \dots, \tau_m$ both generate $\mathbb{R}[V]^G$. Then for all i and j we have $\sigma_i = p_i(\tau_1, \dots, \tau_m)$ and $\tau_j = q_j(\sigma_1, \dots, \sigma_n)$ for polynomials p_i and q_j. If $c^\sigma = (c_1, \dots, c_n)$ is a curve in

$\sigma(V)$, then $c^\tau = (q_1(c^\sigma), \dots, q_m(c^\sigma))$ defines a curve in $\tau(V)$ of the same regularity. Any lift $\bar{c}$ to V of the curve c^σ, i.e., $c^\sigma = \sigma \circ \bar{c}$, is a lift of c^τ as well (and conversely):

$$c^\tau = (q_1(c^\sigma), \dots, q_m(c^\sigma)) = (q_1(\sigma(\bar{c})), \dots, q_m(\sigma(\bar{c}))) = (\tau_1(\bar{c}), \dots, \tau_m(\bar{c})) = \tau \circ \bar{c}.$$

CHAPTER 9

Isometric action of Lie groups and invariants

9.1. Differentiable invariants

We will prove a stronger variant of Schwarz's theorem 8.2.5 which is due to Mather [**Mat77**]. This approach is simpler than Schwarz's original proof.

9.1.1. THEOREM. *Consider a finite dimensional representation $\rho : G \to \mathrm{O}(V)$ of a compact Lie group G. Let $\sigma_1, \ldots, \sigma_n$ be generators for the algebra $\mathbb{R}[V]^G$ of G-invariant polynomials on V. If $\sigma := (\sigma_1, \ldots, \sigma_n) : V \to \mathbb{R}^n$, then*

$$\sigma^* : C^\infty(\mathbb{R}^n) \to C^\infty(V)^G$$

is split surjective, i.e., there is a continuous $\mathbb{R}$-linear right inverse.

9.1.2. Smooth structure on orbit spaces. Let M be a smooth manifold and let G be a compact Lie group acting on M. A function $f : M/G \to \mathbb{R}$ will be called smooth if $f \circ \pi$ is smooth, where $\pi : M \to M/G$ is the projection. The $\mathbb{R}$-algebra $C^\infty(M/G)$ of smooth functions on M/G is then isomorphic to $C^\infty(M)^G$ via π^*. We provide $C^\infty(M)^G$ with the C^∞ topology, and topologize $C^\infty(M/G)$ such that π^* is a homeomorphism. For $x \in M/G$ let I_x denote the ideal in $C^\infty(M/G)$ consisting of functions vanishing at x. The Zariski tangent space to M/G at x is defined as $T_x(M/G) = (I_x/I_x^2)^*$. Here V^* means the vector space of continuous linear real-valued functions on V, and I_x/I_x^2 is provided with the induced topology. A mapping $f : M/G \to N/H$ between orbit spaces is said to be smooth if $f^*C^\infty(N/H) \subseteq C^\infty(M/G)$. For any $x \in M/G$ we define the derivative $df(x) : T_x(M/G) \to T_{f(x)}(N/H)$ as the dual of $f^* : I_{f(x)}/I_{f(x)}^2 \to I_x/I_x^2$. These definitions apply in the particular case when $N/H = N$.

9.1.3. LEMMA. *Consider a finite dimensional representation $\rho : G \to \mathrm{O}(V)$ of a compact Lie group G. Let $\sigma_1, \ldots, \sigma_n$ be a minimal system of generators for $\mathbb{R}[V]^G$. Let $\sigma = (\sigma_1, \ldots, \sigma_n) : V \to \mathbb{R}^n$ and let $\bar\sigma : V/G \to \mathbb{R}^n$ denote the induced mapping. Then $d\bar\sigma(0) : T_0(V/G) \to T_0\mathbb{R}^n$ is an isomorphism.*

PROOF. Denote by Σ, respectively Π, the ideal in $\mathbb{R}[V]$, respectively $\mathbb{R}[\mathbb{R}^n]$, consisting of all polynomials vanishing at 0. Let I, respectively J, denote the ideal in $C^\infty(V)$, respectively in $C^\infty(\mathbb{R}^n)$, consisting of all functions which vanish at 0. By $\hat I$, respectively $\hat J$, we denote the maximal ideal in $\mathbb{R}[[V]]$, respectively $\mathbb{R}[[\mathbb{R}^n]]$, the formal power series in V, respectively $\mathbb{R}^n$. We have a commuting diagram

$$\begin{tikzcd}
\hat J/\hat J^2 \arrow[r] & \hat I^G/(\hat I^G)^2 \\
J/J^2 \arrow[r] \arrow[u] & I^G/(I^G)^2 \arrow[u] \\
\Pi/\Pi^2 \arrow[r] \arrow[u] & \Sigma^G/(\Sigma^G)^2 \arrow[u]
\end{tikzcd}$$

where the horizontal arrows are induced by σ^* and the vertical arrows are induced by the inclusion of the polynomials in the smooth functions, and the Taylor homomorphism, respectively.

We claim that the bottom horizontal arrow is an isomorphism. It suffices to prove that it is injective. If not, we would have a linear combination $\sum_i a_i\sigma_i$ of the σ_i with not all coefficients equal to zero, such that $\sum_i a_i\sigma_i \in (\Sigma^G)^2$. Suppose $a_i \neq 0$. Then

$$\sigma_i = \sum_{j\neq i} b_j\sigma_j + \tau, \quad b_j \in \mathbb{R}, \quad \tau \in (\Sigma^G)^2.$$

If we think of the right side as a sum of monomials in the σ_j, and drop all monomials of degree different from the degree of σ_i, we obtain an expression of σ_i as a polynomial in σ_j, $j \neq i$, which contradicts the hypothesis that $\sigma_1, \dots, \sigma_n$ is minimal. This establishes the claim.

We claim that $\Sigma^G \cap \Sigma^{d+1} \subseteq (\Sigma^G)^2$, where d is the maximum of the degrees of the σ_i. For: Let $u \in \Sigma^G \cap \Sigma^{d+1}$ be homogeneous, and write it as a sum of monomials in the σ_i, all of the same degree as u. Since any such monomial must be a product of two or more σ_i, it follows that $u \in (\Sigma^G)^2$. Since any element of $\Sigma^G \cap \Sigma^{d+1}$ is a sum of homogeneous elements, we obtain the desired inclusion.

Using these two claims, it is easy to check that the top horizontal arrow in the forgoing diagram is an isomorphism. The left vertical arrows are isomorphisms. The composition of the right vertical arrows is an isomorphism, and the kernel of the vertical upper right corner is the closure of 0. Thus we see that the induced mapping

$$\frac{J}{J^2} \to \frac{I^G/(I^G)^2}{\bar{0}}$$

is an isomorphism. The lemma follows then immediately. □

Note that the $\mathbb{R}$-algebra $\mathbb{R}[[\mathbb{R}^n]]$ of formal power series in the coordinates of $\mathbb{R}^n$ is topologized by identifying $\mathbb{R}[[\mathbb{R}^n]]$ with $\mathbb{R}^\infty$ assigning a formal power series the collection of its coefficients. Then $\mathbb{R}^\infty$ is given the product topology.

The mapping $\sigma^* : \mathbb{R}[[\mathbb{R}^n]] \to \mathbb{R}[[V]]^G$ is defined by $\sigma^* f = f \circ \sigma$, where the latter means the power series obtained from f by substituting σ_i for the i-th coordinate of $\mathbb{R}^n$.

9.1.4. Embeddings. We say that $f : M/G \to N$ is an *embedding*, if it is smooth, proper, injective, and $df(x)$ is injective for all $x \in M/G$.

THEOREM. *If $f : M/G \to N$ is proper and smooth, and $f^*C^\infty(N)$ is dense in $C^\infty(M/G)$, then f is an embedding.*

PROOF. We claim that $C^\infty(M/G)$ separates points in M/G. For, let x and y be distinct points in M/G. Since $\pi^{-1}(x)$ and $\pi^{-1}(y)$ are disjoint closed subsets of M, there exists a smooth function u on M such that $u|_{\pi^{-1}(x)} = 0$ and $u|_{\pi^{-1}(y)} = 1$. By averaging over G, we may assume u is invariant. This provides a function $\bar{u} \in C^\infty(M/G)$ with $\bar{u}(x) = 0$ and $\bar{u}(y) = 1$.

Since $C^\infty(M/G)$ separates points, and $f^*C^\infty(N)$ is dense in $C^\infty(M/G)$, it follows that $f^*C^\infty(N)$ separates points. Hence f is injective.

Let $x \in M/G$. Clearly, $f^*(I_{f(x)}/I^2_{f(x)})$ is dense in I_x/I_x^2. Consequently, the induced mapping

$$\frac{I_{f(x)}}{I^2_{f(x)}} \to \frac{I_x/I_x^2}{\bar{0}}$$

is surjective, since the vector space on the right is finite dimensional and Hausdorff; $\bar{0}$ denotes the closure of 0. Passing to the duals, we obtain that $df(x) : T_x(M/G) \to T_{f(x)}N$ is injective. □

9.1.5. $\sigma' : V/G \to \mathbb{R}^n$ is an embedding. We have seen in lemma 8.2.2 that the orbit map σ is proper. The polynomials $\mathbb{R}[V]$ are dense in $C^\infty(V)$. Thus, by averaging over G, we see that $\mathbb{R}[V]^G$ is dense in $C^\infty(V)^G$. Since $\sigma^* : \mathbb{R}[\mathbb{R}^n] \to \mathbb{R}[V]^G$ is surjective (Hilbert's theorem), we find that $\sigma^*\mathbb{R}[\mathbb{R}^n]$ is dense in $C^\infty(V)^G$. Hence $\sigma^* C^\infty(\mathbb{R}^n)$ is dense in $C^\infty(V)^G$.

That shows, by theorem 9.1.4, that the induced map $\sigma' : V/G \to \mathbb{R}^n$ is an embedding. Theorem 9.1.1 will then follow from the following theorem.

9.1.6. THEOREM. *If $f : M/G \to N$ is an embedding, then f^* is split surjective.*

Let us fix some notation for convenience: We denote by R_n the assertion that theorem 9.1.1 holds for all orthogonal actions on $\mathbb{R}^p$, where $p \le n$. Let T_n be the assertion that theorem 9.1.6 holds for all smooth actions on manifolds of dimension less or equal than n, and all embeddings of the associated quotient spaces.

9.1.7. Local split surjectivity implies global split surjectivity. Suppose $f : M/G \to N$ is a smooth embedding. Let $x \in M/G$. We will say that f^* is *locally split surjective at x* if there is an open neighborhood U of x in M/G and a continuous linear mapping $l : C^\infty(M/G) \to C^\infty(N)$ such that $f^*l(u)|_U = u|_U$ for any $u \in C^\infty(M/G)$. We will say that f^* is *locally split surjective* if it is locally split surjective at each point of M/G.

LEMMA. *If f^* is locally split surjective, then it is split surjective.*

PROOF. Since $f(M/G)$ is closed in N and N is paracompact, we can choose a locally finite open cover $\{W_\alpha\}$ of N such that for each α, there is a continuous linear mapping $l_\alpha : C^\infty(M/G) \to C^\infty(N)$ such that $f^*l_\alpha(u)|_{f^{-1}(W_\alpha)} = u|_{f^{-1}(W_\alpha)}$, for any $u \in C^\infty(M/G)$. Let $\{\rho_\alpha\}$ be a smooth partition of unity subordinate to $\{W_\alpha\}$. We define $l : C^\infty(M/G) \to C^\infty(N)$ by $l(u) = \sum_\alpha \rho_\alpha l_\alpha(u)$. It is easily verified that l is continuous and $f^*l = \mathrm{id}$. □

9.1.8. R_n implies T_n. We may construct a G-invariant metric on M by taking any Riemannian metric and averaging over G, see theorem 9.4.9. Let $x \in M$, let $G.x$ be the orbit through x, and let G_x denote the isotropy group at x. Denote by E_x the vector space of tangent vectors at x, perpendicular to $G.x$. It follows from the differentiable slice theorem 9.4.4 that there exists a diffeomorphism $\varphi : U_x/G \to E_x^\epsilon/G_x$ (i.e. a smooth mapping with a smooth inverse in the sense defined above), where U_x is an open invariant neighborhood of $G.x$ in M and E_x^ϵ is the set of vectors in E_x of norm less than $\epsilon > 0$.

Since E_x is an Euclidean space and G_x acts orthogonally on it (see definition 9.5.2), we may choose a minimal Hilbert basis $\sigma_1, \dots, \sigma_k$ of $\mathbb{R}[E_x]^{G_x}$. Let us assume that R_n holds. Hence $\sigma^* : C^\infty(\mathbb{R}^k) \to C^\infty(E_x)^{G_x}$ is split surjective, and, therefore, $\sigma^* : C^\infty(\mathbb{R}^k) \to C^\infty(E_x^\epsilon)^{G_x}$ is locally split surjective at 0.

Let $y = f(\bar{x})$, where $\bar{x}$ is the image of x under the projection $\pi : M \to M/G$. Let $y_1, \dots, y_p$ be a smooth local system of coordinates on N, defined in a neighborhood of y. Since $\sigma^* : C^\infty(\mathbb{R}^k) \to C^\infty(E_x^\epsilon)^{G_x}$ is locally split surjective at 0, there exist smooth function $u_1, \dots, u_p$, defined in a neighborhood of 0 in $\mathbb{R}^k$, such that $\sigma^* u_i = y_i \circ f \circ \varphi^{-1}$ in a sufficiently small neighborhood of 0 in E_x^ϵ/G_x.

Let $u : \mathbb{R}^k \to N$ be defined by $y_i \circ u = u_i$; then u is defined in a neighborhood of 0 and we have the following commuting diagram

$$\begin{tikzcd} U_x/G \arrow[r, "f"] \arrow[d, "\varphi"'] & N \\ E_x^\epsilon/G_x \arrow[r, "\bar{\sigma}"] & \mathbb{R}^k \arrow[u, "u"'] \end{tikzcd}$$

We find that $d\varphi(\bar{x})$ is an isomorphism, since φ is a diffeomorphism. Moreover, $d\bar{\sigma}(0)$ is an isomorphism, by lemma 9.1.3. Since f is an embedding, $df(\bar{x})$ is injective. So $du(0)$ is injective, because $df(\bar{x}) = du(0) \cdot d\bar{\sigma}(0) \cdot d\varphi(\bar{x})$. Thus $u^* : C^\infty(N) \to C^\infty(\mathbb{R}^k)$ is locally split surjective at 0.

Now $f^* = \varphi^* \bar{\sigma}^* u^*$ is a composition of locally split surjective homomorphisms. Hence f^* is locally split surjective at $\bar{x}$. Since $\bar{x}$ is an arbitrary point of M/G, we may conclude, by lemma 9.1.7, that f^* is split surjective.

Thus we have showed that R_n implies T_n.

9.1.9. A variant of Borel's lemma. Let $\mathrm{Tay} : C^\infty(\mathbb{R}^n) \to \mathbb{R}[[\mathbb{R}^n]]$ be the Taylor homomorphism which associates to $f \in C^\infty(\mathbb{R}^n)$ its Taylor series expansion $\mathrm{Tay}(f)$ at 0. A lemma due to E. Borel states that Tay is surjective. Unfortunately, Tay is not split surjective.

Therefore, we investigate the following lifting problem. Suppose $F : C^\infty(\mathbb{R}^k) \to \mathbb{R}[[\mathbb{R}^n]]$ is a continuous linear mapping. Find a continuous linear mapping $\tilde{F} : C^\infty(\mathbb{R}^k) \to C^\infty(\mathbb{R}^n)$ such that the following diagram commutes

$$\begin{tikzcd} C^\infty(\mathbb{R}^k) \arrow[d, dotted, "\tilde{F}"'] \arrow[dr, "F"] & \\ C^\infty(\mathbb{R}^n) \arrow[r, "\mathrm{Tay}"] & \mathbb{R}[[\mathbb{R}^n]] \end{tikzcd} \tag{9.1}$$

Let $x \in \mathbb{R}^k$. We will say that F *is null at* x if there exists a neighborhood U of x in $\mathbb{R}^k$ such that if $f \in C^\infty(\mathbb{R}^k)$ and $\mathrm{supp}(f) \subseteq U$, then $F(f) = 0$. Clearly, the set of points at which F is null is open. By the *support of* F, denoted by $\mathrm{supp}(F)$, we mean the complement of the set of points where F is null. This definition of support generalizes the standard definition of the support of a distribution.

LEMMA. *If F has compact support, then there exists a continuous linear $\tilde{F}$ which makes diagram* (9.1) *commute.*

PROOF. Let p be a smooth function on $\mathbb{R}^k$ with compact support and values in $[0,1]$ such that p is identically 1 in a neighborhood of $\mathrm{supp}(F)$. Let ρ be a smooth function on $\mathbb{R}^n$ with support in the unit ball, invariant under the action of the orthogonal group on $\mathbb{R}^n$, and identically 1 in a neighborhood of 0. For any $\lambda > 0$, let ρ_λ be the function on $\mathbb{R}^n$ defined by $\rho_\lambda(x) = \rho(\lambda x)$.

Let K be a large positive number, such that p has support in the interior of the cube D_K of side K centered at 0 in $\mathbb{R}^k$. For any $\alpha = (\alpha_1, \ldots, \alpha_k) \in \mathbb{Z}^k$, let $e_\alpha(x) = e^{2\pi i \langle \alpha | x \rangle / K}$, if $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$. Then $e_\alpha \in C^\infty(\mathbb{R}^k)$. Let $\epsilon_\alpha = F(e_\alpha)$ and let $\epsilon_{\alpha,r}$ be the homogeneous part of order r of ϵ_α. Then $\epsilon_\alpha = \sum_{r \geq 0} \epsilon_{\alpha,r}$.

We will choose later, for each $\alpha \in \mathbb{Z}^k$ and each $r \geq 0$, a positive number $\lambda(\alpha, r)$. If $u \in C^\infty(\mathbb{R}^k)$, we expand $pu|_{D_K}$ in a Fourier series

$$pu|_{D_K} = \sum_\alpha c_\alpha e_\alpha, \quad c_\alpha \in \mathbb{C}.$$

Then we define

$$\tilde{F}(u) = \sum_{\alpha, r} c_\alpha \rho_{\lambda(\alpha,r)} \epsilon_{\alpha,r}, \tag{9.2}$$

where we think of $\epsilon_{\alpha,r}$ as a function on $\mathbb{R}^n$, which we may, since it is a homogeneous polynomial of degree r.

We will show that by choosing the $\lambda(\alpha, r)$ suitably, we may arrange that the sum on the right-hand side of (9.2) converges with respect to the C^∞ topology on $C^\infty(\mathbb{R}^n)$, and that the resulting mapping $\tilde{F} : C^\infty(\mathbb{R}^k) \to C^\infty(\mathbb{R}^n)$ is continuous. Obviously, $\tilde{F}$ is linear, if it can be defined in this way, and $\mathrm{Tay} \circ \tilde{F} = F$.

We shall also show that we may choose the $\lambda(\alpha, r)$ so that $\lambda(\alpha, r) = \lambda(\alpha', r)$ if $|\alpha_i| = |\alpha_i'|$, for all $1 \le i \le k$. Then, if u is real-valued, so is $\tilde{F}(u)$.

First, let us estimate the size of $\epsilon_{\alpha,r}$. Let $\mathbb{R}[\mathbb{R}^n]_r$ denote the vector space of all homogeneous polynomials on $\mathbb{R}^n$ of degree r. We let $\|.\|$ denote any norm on $\mathbb{R}[\mathbb{R}^n]_r$ (since $\mathbb{R}[\mathbb{R}^n]_r$ is finite dimensional, any two norms are equivalent). Define $F_r : C^\infty(\mathbb{R}^k) \to \mathbb{R}[\mathbb{R}^n]_r$ by $F_r(u) = F(u)_r$, where the latter denotes the homogeneous part of order r of $F(u)$. Since F is linear, continuous, and has compact support, it follows that F_r has the same properties. Thus, F_r is a $\mathbb{R}[\mathbb{R}^n]_r$-valued distribution on $\mathbb{R}^k$ with compact support. It follows from a standard estimate in the theory of distributions (cf. [**Hör83a**, (1.5.4)]), that

$$\|F_r(u)\| \le C_r \sum_{0 \le |\beta| \le s(r)} \sup_{D_K} \|\partial^\beta u\|, \quad u \in C^\infty(\mathbb{R}^k),$$

where C_r and $s(r)$ are constants, and $\partial^\beta = (\partial_1)^{\beta_1} \cdots (\partial_k)^{\beta_k}$. From this, we obtain immediately

$$\|\epsilon_{\alpha,r}\| = \|F(e_\alpha)_r\| \le C_r'(1 + |\alpha|)^{s(r)}, \tag{9.3}$$

where C_r' is a constant and $|\alpha| = \sum_{j=1}^k |\alpha_j|$.

Now we estimate the l-th total derivative of $\tilde{F}(u)$. From (9.2) and Leibniz's formula, we get

$$\tilde{F}(u)^{(l)} = \sum_{\alpha,r} c_\alpha \sum_{0 \le m \le l} \binom{l}{m} \rho^{(m)}_{\lambda(\alpha,r)} \epsilon^{(l-m)}_{\alpha,r}.$$

Clearly, $\rho^{(m)}_{\lambda(\alpha,r)}$ has support in the ball of radius $\lambda(\alpha, r)^{-1}$, and $\sup \|\rho^{(m)}_{\lambda(\alpha,r)}\| \le C_m'' \lambda(\alpha, r)^m$, where $C_m'' = \sup \|\rho^{(m)}\|$. If A denotes the ball with radius $\lambda(\alpha, r)^{-1}$, we find, in view of (9.3), and the fact that $\epsilon_{\alpha,r}$ is a homogeneous polynomial of degree r, that

$$\sup_A \|\epsilon^{(l-m)}_{\alpha,r}\| \le C_r'''(1 + |\alpha|)^{s(r)} \lambda(\alpha, r)^{l-m-r},$$

where C_r''' is a constant. Thus,

$$\begin{aligned} \sup \|\tilde{F}(u)^{(l)}\| &\le \sum_{\alpha,r} c_\alpha \sum_{0 \le m \le l} \binom{l}{m} C_m'' C_r'''(1 + |\alpha|)^{s(r)} \lambda(\alpha, r)^{l-r} \\ &\le \sum_{\alpha,r} c_\alpha C^{iv}_{rl}(1 + |\alpha|)^{s(r)} \lambda(\alpha, r)^{l-r}, \end{aligned} \tag{9.4}$$

where

$$C^{iv}_{rl} = \sum_{0 \le m \le l} \binom{l}{m} C_m'' C_r'''.$$

Now we choose the $\lambda(\alpha, r)$. For each fixed r, we will choose $\lambda(\alpha, r)$ to be so rapidly increasing with $|\alpha|$ that

$$\sum_\alpha C^{iv}_{rl}(1 + |\alpha|)^{s(r)} \lambda(\alpha, r)^{l-r} \le 2^{l-r}, \quad \text{if } l < r, \tag{9.5}$$

and at the same time, choose $\lambda(\alpha, r)$ so that for fixed r it has at most polynomial growth in $|\alpha|$. This can be achieved, for example, by choosing

$$\lambda(\alpha, r) = C^v_r(1 + |\alpha|)^{s(r)+k+1}, \tag{9.6}$$

where

$$C^v_r = \sup_{l<r} 2 \left(C^{iv}_{rl} \sum_\alpha (1 + |\alpha|)^{-k-1} \right)^{\frac{1}{r-l}}.$$

Since α varies over $\mathbb{Z}^k$, the sum $\sum_\alpha (1 + |\alpha|)^{-\varrho}$ converges for any $\varrho > k$. Since there are only finitely many $l < r$, it is then clear that $C^v_r < \infty$. Evidently, $\lambda(\alpha, r)$

satisfies (9.5) and has at most polynomial growth in $|\alpha|$. Moreover, $|\alpha_i| = |\alpha'_i|$, for $1 \leq i \leq k$, implies obviously $\lambda(\alpha, r) = \lambda(\alpha', r)$.

Let us show that (9.2) converges in the C^∞ topology. Since the c_α are the Fourier coefficients of $pu|_{D_K}$, and pu is C^∞ and vanishes in a neighborhood of ∂D_K, we have for any $\mu > 0$,

$$(1 + |\alpha|)^\mu c_\alpha \to 0, \quad \text{as } |\alpha| \to \infty. \tag{9.7}$$

From (9.4), (9.5), and (9.6), we obtain

$$\sup \|\tilde{F}(u)^{(l)}\| \leq \sum_{0 \leq r \leq l} C^{vi}_{rl} \sum_\alpha c_\alpha (1 + |\alpha|)^{s'(r,l)} + \sum_{r>l} 2^{l-r} \sum_\alpha c_\alpha, \tag{9.8}$$

where

$$C^{vi}_{rl} = C^{iv}_{rl} (C^v_r)^{l-r}, \quad s'(r,l) = s(r) + (l-r)(s(r) + k + 1).$$

From (9.7) it clearly follows that (9.8) converges. Moreover, if we define, for $u \in C^\infty(\mathbb{R}^k)$, $pu|_{D_K} = \sum c_\alpha e_\alpha$,

$$\|u\|_\mu = \sum_\alpha |c_\alpha| (1 + |\alpha|)^\mu,$$

then $\|.\|_\mu$ is a continuous semi-norm on $C^\infty(\mathbb{R}^k)$. By (9.8), we find

$$\sup \|\tilde{F}(u)^{(l)}\| \leq \sum_{0 \leq r \leq l} C^{vi}_{rl} \|u\|_{s'(r,l)} + C\|u\|_0.$$

Therefore $\tilde{F}$ is continuous with respect to the C^∞-topologies on $C^\infty(\mathbb{R}^k)$ and $C^\infty(\mathbb{R}^n)$. This completes the proof of the lemma. □

9.1.10. T_{n-1} implies that $\sigma^* : C^\infty(\mathbb{R}^n)_0 \to C^\infty(V)^G_0$ is split surjective. Let us suppose that T_{n-1} holds, i.e., theorem 9.1.6 holds for all smooth actions on manifolds of dimension less or equal than $n-1$, and all embeddings of the associated quotient spaces.

We consider an orthogonal action of a compact Lie group G on a finite dimensional Euclidean vector space V. Let $\sigma_1, \dots, \sigma_n$ be a minimal system of homogeneous generators of $\mathbb{R}[V]^G$ with degrees $d_1, \dots, d_n$, respectively. Let $C^\infty(\mathbb{R}^n)_0$ denote the set of C^∞-functions on $\mathbb{R}^n$ which vanish of infinite order at 0. We will show that $\sigma^* : C^\infty(\mathbb{R}^n)_0 \to C^\infty(V)^G_0$ is split surjective.

Consider the unit sphere S^{n-1} in $\mathbb{R}^n$ and set $\Sigma^{n-1} = \sigma^{-1}(S^{n-1})$. Each ray emanating from the origin in V meets Σ^{n-1} in exactly one point, transversally. For: Consider $x \in V \backslash \{0\}$ and let $\varphi(t) = |\sigma(tx)|^2$ for $t \geq 0$. Then Σ^{n-1} meets the ray through x in exactly those points where $\varphi(t) = 1$. But, $\varphi(t) = t^{2d_1}\sigma_1^2(x) + \cdots + t^{2d_n}\sigma_n^2(x)$ and not all $\sigma_i(x)$ are 0, since $\langle x \mid x \rangle$ is an invariant and therefore can be expressed as a polynomial in the σ_i. It follows that $\varphi(t) = 1$ has exactly one positive solution, and $\varphi'(t) \neq 0$ there, which proves that the ray through x meets Σ^{n-1} in exactly one point, transversally.

Hence Σ^{n-1} is a compact analytic manifold.

We define a mapping $r : S^{n-1} \times \mathbb{R} \to \mathbb{R}^n$ by $r(x_1, \dots, x_n, t) = (t^{d_1}x_1, \dots, t^{d_n}x_n)$ and a mapping $\rho : \Sigma^{n-1} \times \mathbb{R} \to V$ by $\rho(x, t) = tx$. Denote by $\mathbb{R}_+$ the non-negative real numbers. Let $r_+ : S^{n-1} \times \mathbb{R}_+ \to \mathbb{R}^n$ and $\rho_+ : \Sigma^{n-1} \times \mathbb{R}_+ \to V$ denote the restrictions of r and ρ.

We consider Σ^{n-1} as a G-space with respect to the restriction of the given action on V. We consider $\mathbb{R}_+$ as a G-space with the trivial action. Then $\Sigma^{n-1} \times \mathbb{R}_+$ is a

G-space and ρ_+ is equivariant. Moreover the following diagram commutes:

$$\begin{tikzcd}
S^{n-1} \times \mathbb{R}_+ \arrow[d, "r_+"'] & \Sigma^{n-1} \times \mathbb{R}_+ \arrow[l, "\sigma\times\mathrm{id}"'] \arrow[d, "\rho_+"] \\
\mathbb{R}^n & V \arrow[l, "\sigma"]
\end{tikzcd}$$

If M is any manifold and K is a subset of M, let $C^\infty(M)_K$ denote the C^∞ functions on M which vanish of infinite order on K.

Lemma. *We have:*

(1) $r_+^* C^\infty(\mathbb{R}^n)_0 = C^\infty(S^{n-1} \times R_+)_{S^{n-1}\times\{0\}}$.
(2) $\rho_+^* C^\infty(V)_0^G = C^\infty(\Sigma^{n-1} \times R_+)^G_{\Sigma^{n-1}\times\{0\}}$.

Proof. In both cases the inclusion '$\subseteq$' is obvious. Moreover, it is clear that the left side is dense in the right, since it contains those functions in the right which vanish in a neighborhood of $S^{n-1} \times \{0\}$, respectively $\Sigma^{n-1} \times \{0\}$. By applying Glaeser's theorem [**Gla63a**], we see that the left side is closed in the right side. (Strictly speaking, one should apply Glaeser's theorem to the mappings r and ρ, and then one finds that $r^* C^\infty(\mathbb{R}^n)$ and $\rho^* C^\infty(V)$ are closed. But, what we need then follows quickly.) $\square$

By the above lemma, in order to show that $\sigma^* : C^\infty(\mathbb{R}^n)_0 \to C^\infty(V)_0^G$ is split surjective, it is enough to prove that the top arrow in the following commutative diagram is split surjective:

$$\begin{tikzcd}
C^\infty(S^{n-1} \times \mathbb{R}_+)_{S^{n-1}\times\{0\}} \arrow[r, "(\sigma\times\mathrm{id})^*"] & C^\infty(\Sigma^{n-1} \times \mathbb{R}_+)^G_{\Sigma^{n-1}\times\{0\}} \\
C^\infty(\mathbb{R}^n)_0 \arrow[u, "r_+^*"] \arrow[r, "\sigma^*"'] & C^\infty(V)_0^G \arrow[u, "\rho_+^*"']
\end{tikzcd}$$

For, it is an easy consequence of the previous lemma that the vertical arrows are homeomorphisms.

Now if U is an open set in $\mathbb{R}^n$ and F is a Fréchet space, we let $C^\infty(U, F)$ denote the set of C^∞ functions on U with values in F, provided with the C^∞ topology. We have a commutative diagram

$$\begin{tikzcd}
C^\infty(S^{n-1} \times \mathbb{R}_+)_{S^{n-1}\times\{0\}} \arrow[r, "(\sigma\times\mathrm{id})^*"] & C^\infty(\Sigma^{n-1} \times \mathbb{R}_+)^G_{\Sigma^{n-1}\times\{0\}} \\
C^\infty(\mathbb{R}_+, C^\infty(S^{n-1}))_0 \arrow[u, "\cong"] \arrow[r] & C^\infty(\mathbb{R}_+, C^\infty(\Sigma^{n-1})^G)_0 \arrow[u, "\cong"']
\end{tikzcd}$$

where the bottom arrow is induced from the mapping $\sigma^* : C^\infty(S^{n-1}) \to C^\infty(\Sigma^{n-1})^G$. But σ^* has dense image by the argument used in 9.1.5, so it is split surjective by theorem 9.1.4 and by the hypothesis that T_{n-1} holds. Therefore, the bottom arrow in the above diagram is split surjective, and we deduce that $(\sigma \times \mathrm{id})^*$ is split surjective.

So we have proved that $\sigma^* : C^\infty(\mathbb{R}^n)_0 \to C^\infty(V)_0^G$ is split surjective, if T_{n-1} is assumed.

9.1.11. End of proof of theorem 9.1.1. We shall show that T_{n-1} if and only if R_n. We have already seen that T_n if and only if R_n in 9.1.5 and 9.1.8. Clearly, T_0 holds, so this will suffice.

Consider an orthogonal action of G on V and let $\sigma_1, \ldots, \sigma_n$ constitute a minimal system of homogeneous generators of $\mathbb{R}[V]^G$. Consider the following commutative diagram

$$\begin{tikzcd} 0 \arrow[r] & C^\infty(V)_0^G \arrow[r] & C^\infty(V)^G \arrow[r, "\mathrm{Tay}"] & \mathbb{R}[[V]]^G \arrow[r] & 0 \\ 0 \arrow[r] & C^\infty(\mathbb{R}^n)_0 \arrow[r] \arrow[u] & C^\infty(\mathbb{R}^n) \arrow[r, "\mathrm{Tay}"] \arrow[u, "\sigma^*"] & \mathbb{R}[[\mathbb{R}^n]] \arrow[r] \arrow[u] & 0 \end{tikzcd} \tag{9.9}$$

We claim that the right vertical arrow is split surjective. For: Let $\mathbb{R}[V]_d^G$ denote the vector space of homogeneous invariant polynomials on V of degree d. Let $x_1, \ldots, x_n$ be the coordinates of $\mathbb{R}^n$ and assign x_i weight $d_i = \deg \sigma_i$. Let $\mathbb{R}[\mathbb{R}^n]_d^w$ be the polynomials on $\mathbb{R}^n$ which are weighted homogeneous of degree d. Hilbert's theorem implies that $\sigma_d^* : \mathbb{R}[\mathbb{R}^n]_d^w \to \mathbb{R}[V]_d^G$ is surjective. Since $\mathbb{R}[\mathbb{R}^n]_d^w$ and $\mathbb{R}[V]_d^G$ are finite dimensional vector spaces, it follows that σ_d^* is split surjective. But $\sigma^* = \bigoplus_d \sigma_d^* : \bigoplus_d \mathbb{R}[\mathbb{R}^n]_d^w \to \bigoplus_d \mathbb{R}[V]_d^G$, so it follows that σ^* is split surjective. Thus the claim is proved.

Let us consider the composition

$$C^\infty(V) \xrightarrow{A} C^\infty(V)^G \xrightarrow{\mathrm{Tay}} \mathbb{R}[[V]]^G \xrightarrow{S} \mathbb{R}[[\mathbb{R}^n]],$$

where A is defined by averaging over G, and S splits $\mathbb{R}[[\mathbb{R}^n]] \to \mathbb{R}[[V]]^G$. By lemma 9.1.9, we may lift this mapping to a continuous linear mapping $\eta_0 : C^\infty(V) \to C^\infty(\mathbb{R}^n)$. Let $\eta_1 = \eta_0|_{C^\infty(V)^G}$. Then we have (in view of diagram (9.9))

$$\mathrm{Tay} \circ \sigma^* \circ \eta_1 = \mathrm{Tay} : C^\infty(V)^G \to \mathbb{R}[[V]]^G.$$

Since the inclusion mapping of $C^\infty(V)_0^G$ into $C^\infty(V)^G$ is a homeomorphism onto its image, it therefore follows that $\sigma^* \circ \eta_1 - \mathrm{id}|_{C^\infty(V)^G}$ may be regarded as a continuous linear mapping into $C^\infty(V)_0^G$.

From the split surjectivity of the left arrow in diagram (9.9) which has been proved in 9.1.10, we can therefore deduce that there exists a continuous linear mapping $\eta_2 : C^\infty(V)^G \to C^\infty(\mathbb{R}^n)_0 \subseteq C^\infty(\mathbb{R}^n)$ such that $\sigma^* \circ \eta_1 - \mathrm{id}|_{C^\infty(V)^G} = \sigma^* \circ \eta_2$. Setting $\eta = \eta_1 - \eta_2 : C^\infty(V)^G \to C^\infty(\mathbb{R}^n)$, we see that $\sigma^* \circ \eta = \mathrm{id}|_{C^\infty(V)^G}$, so σ^* is split surjective.

This completes the proof of theorem 9.1.1.

9.2. The space $\sigma(V)$

Remember the characterization of the space of hyperbolic polynomials with a fixed degree, given in theorem 3.1.2. There is a similar description of the orbit space of an arbitrary finite dimensional orthogonal representation of a compact Lie group.

9.2.1. The generalized Bezoutiant. We work in the setting presented in section 8.2. Let $\langle . \mid . \rangle$ denote also the G-invariant dual inner product on the dual space V^*. The differentials $d\sigma_i : V \to V^*$ are G-equivariant:

$$l_g^* d\sigma_i(v)(z) = d\sigma_i(g.v)(dl_g(v)(z)) = d(\sigma_i \circ l_g)(v)(z) = d\sigma_i(g.v)(z)$$

for arbitrary v and z, where $l_g : V \to V$ denotes the left-action by the element $g \in G$. Therefore, the polynomials $v \mapsto \langle d\sigma_i(v) \mid d\sigma_j(v) \rangle$ are in $\mathbb{R}[V]^G$, and they are entries of an $n \times n$ symmetric matrix valued polynomial

$$B(v) := \begin{pmatrix} \langle d\sigma_1(v) \mid d\sigma_1(v) \rangle & \ldots & \langle d\sigma_1(v) \mid d\sigma_n(v) \rangle \\ \vdots & \ddots & \vdots \\ \langle d\sigma_n(v) \mid d\sigma_1(v) \rangle & \ldots & \langle d\sigma_n(v) \mid d\sigma_n(v) \rangle \end{pmatrix}. \tag{9.10}$$

There is a unique matrix valued polynomial $\tilde{B}$ on $V/\!\!/G$ such that $B = \tilde{B} \circ \sigma$.

Note that in the particular case of hyperbolic polynomials this matrix B reduces to the Bezoutiant defined in section 3.1: Then $G = \mathrm{S}_n$ acts on $V = \mathbb{R}^n$ by permuting the coordinates, and $\mathbb{R}[V]^G = \mathbb{R}[\sigma_1, \dots, \sigma_n]$, where $\sigma_1, \dots, \sigma_n$ are the elementary symmetric polynomials which are algebraically independent, whence $V/\!\!/G = \mathbb{R}^n$. We may choose different generators $s_1, \frac{1}{2}s_2, \dots, \frac{1}{n}s_n$ of $\mathbb{R}[V]^G$, where the $s_i(x_1, \dots, x_n) = \sum_{j=1}^n x_j^i$ are the Newton polynomials. Then, for $x = (x_1, \dots, x_n)$,

$$\begin{aligned} \langle d(1/is_i)(x) \mid d(1/js_j)(x) \rangle &= \langle (x_1^{i-1}, x_2^{i-1}, \dots, x_n^{i-1}) \mid (x_1^{j-1}, x_2^{j-1}, \dots, x_n^{j-1}) \rangle \\ &= \sum_{k=1}^n x_k^{i+j-2} = s_{i+j-2} \end{aligned}$$

are the entries of the Bezoutiant.

9.2.2. Description of $\sigma(V)$. We have seen in theorem 3.1.2 that for the standard representation of the symmetric group S_n on $\mathbb{R}^n$ and σ_i the elementary symmetric functions we find $\sigma(\mathbb{R}^n) = \{z \in \mathbb{R}^n : \tilde{B}(z) \geq 0\}$, where for a real symmetric matrix A let $A \geq 0$ indicate that A is positive semidefinite. The semialgebraic set $\sigma(\mathbb{R}^n)$ can be identified with Hyp_n, the space of monic hyperbolic polynomials of fixed degree n. (The set $\sigma(V)$ is independent of the choice of generators, see remark 9.2.3(1).)

The following theorem provides a generalization of this special case. We follow the approach of Procesi and Schwarz [**PS85**]. It seems that Abud and Sartori [**AS81**] and [**AS83**] have been the first to realize the description of $\sigma(V)$ given in this theorem.

THEOREM. *Consider a real finite dimensional representation $\rho : G \to \mathrm{O}(V)$ of a compact Lie group G. Let $\sigma_1, \dots, \sigma_n$ be generators for the algebra $\mathbb{R}[V]^G$ of G-invariant polynomials on V. Let $\sigma := (\sigma_1, \dots, \sigma_n) : V \to \mathbb{R}^n$, and let $\tilde{B}$ be the unique matrix valued polynomial on $V/\!\!/G$ such that $B = \tilde{B} \circ \sigma$, where B is defined in* (9.10)*. Then, we have $\sigma(V) = \{z \in V/\!\!/G : \tilde{B}(z) \geq 0\}$.*

The proof of this theorem is displaced into the next section. It requires some background on linear reductive groups which would go beyond the scope of the present section. The following section gives a detailed exposure of the proof of theorem 9.2.2 and of the underlying properties of linear reductive groups.

9.2.3. REMARKS. (1) The sets $\sigma(V)$ and $V/\!\!/G$ and our descriptions of them depend upon our choice of generators for $\mathbb{R}[V]^G$, but not in a serious way: Let $\mathcal{Z}$ denote the variety of real maximal ideals of $\mathbb{R}[V]^G$ and let $\mathcal{X} = \pi(V)$, where $\pi : V \to \mathcal{Z}$ is dual to the inclusion $\mathbb{R}[V]^G \subseteq \mathbb{R}[V]$. Then $V/\!\!/G$ and $\mathcal{Z}$ are canonically isomorphic, and the inequalities defining $\sigma(V)$ as a subset of $V/\!\!/G$, thought of as inequalities involving elements of $\mathbb{R}[\mathcal{Z}] = \mathbb{R}[V]^G$, define $\mathcal{X}$ as a subset of $\mathcal{Z}$. Hence changing the choice of generators may change the inequalities, but not the set they describe.

(2) Choose an orthonormal basis $v_1, \dots, v_m$ of V relative to $\langle . \mid . \rangle$. Then, relative to these coordinates, B is the matrix of inner products of the gradients of the σ_i; equivalently, $B = JJ^\top$, where $J = (\frac{\partial \sigma_i}{\partial v_j})_{ij}$ is the Jacobian matrix of σ. Note that J generalizes the Vandermonde matrix of the symmetric group case (section 3.1).

9.2.4. For each $1 \leq i_1 < \cdots < i_s \leq n$ and $1 \leq j_1 < \cdots < j_s \leq n$, for $s \leq n$, consider the matrix with entries $\langle d\sigma_{i_p} \mid d\sigma_{j_q} \rangle$ for $1 \leq p, q \leq s$, an s-minor of B. Denote its determinant by $\Delta_{i_1,\dots,i_s}^{j_1,\dots,j_s}$. Then, $\Delta_{i_1,\dots,i_s}^{j_1,\dots,j_s}$ is a G-invariant polynomial on V, and thus there is a unique polynomial $\tilde{\Delta}^{j_1,\dots,j_s}_{i_1,\dots,i_s}$ on $V/\!\!/G$ such that $\Delta^{j_1,\dots,j_s}_{i_1,\dots,i_s} = \tilde{\Delta}^{j_1,\dots,j_s}_{i_1,\dots,i_s} \circ \sigma$. Recall from linear algebra that the real symmetric matrix $\tilde{B}(z)$ is positive semidefinite if and only if all its principal (i.e. symmetric) minors $\tilde{\Delta}^{i_1,\dots,i_s}_{i_1,\dots,i_s}(z)$ are non-negative.

9.3. Linear reductive groups

The scope of this section is the proof of theorem 9.2.2. It is based essentially on some results on linear reductive groups due to Kempf and Ness [**KN79**] and Dadok and Kac [**DK85**] and on Luna's slice theorem 9.3.5. We will carefully develop these ingredients and finally prove theorem 9.2.2 at the end of the section.

9.3.1. Rational actions. All algebraic groups, unless otherwise stated, shall be linear and defined over $\mathbb{C}$ throughout this section. Let K be an algebraic group. An action of K on a complex algebraic variety U is said to be *rational*, if the canonical map $l : K \times U \to U$ is a morphism of varieties. We shall use the following notation: $l(k,u) = l_k(u) = l^u(k) = k.u$. For instance, if $K \to \mathrm{GL}(W)$ is a representation of K on a finite dimensional complex vector space W, then W is a rational K-variety.

An algebraic group K is called *reductive*, if it is linear and if every finite dimensional complex representation of K is *completely reducible*, i.e., if the action of K on the representation space leaves invariant some linear subspace, then it leaves invariant a complementary linear subspace also. Note that often an algebraic group is said to be reductive, if its unipotent radical is trivial (e.g. [**Bor69**]). In general this differs from our definition, but in characteristic 0 the two terms are interchangeable (see [**Bor66**]).

Unless otherwise stated, all topological notions in this section shall refer to the Zariski topology.

9.3.2. Reynolds operator. Let K be a reductive algebraic group and U a rational affine K-variety. A *Reynolds operator* is a K-invariant projection, i.e., a linear map $E : \mathbb{C}[U] \to \mathbb{C}[U]^K$ such that:

(*i*) $E(f) = f$ for all $f \in \mathbb{C}[U]^K$.
(*ii*) E is K-invariant, i.e., $E(k.f) = E(f)$ for all $f \in \mathbb{C}[U]$ and all $k \in K$.

Note that for finite groups the Reynolds operator is just averaging.

Lemma. *Let K be a reductive algebraic group and U a rational affine K-variety. Then there exists a unique Reynolds operator $E : \mathbb{C}[U] \to \mathbb{C}[U]^K$. Moreover, E has the following properties:*

(1) *If $S \subseteq \mathbb{C}[U]$ is a K-stable subspace, then $E(S) = S^K$.*
(2) *E is a $\mathbb{C}[U]^K$-module homomorphism, i.e., $E(fg) = fE(g)$ if $f \in \mathbb{C}[U]^K$ and $g \in \mathbb{C}[U]$.*

Proof. It is easy to see that reductiveness of K implies: For any finite dimensional and K-invariant $V \subseteq \mathbb{C}[U]$ we have a decomposition $V = V^K \oplus W$ with a unique subrepresentation W and $(W^*)^K = \{0\}$. We show uniqueness first: If $E : \mathbb{C}[U] \to \mathbb{C}[U]^K$ is any Reynolds operator, then its restriction to V^K must be the indentity and its restriction to W must be 0, since otherwise we would have a non-zero element in $(W^*)^K$. Let us prove existence: We define $E_V : V = V^{\oplus} W \to V^K$ as the projection onto V^K along W. If V' is another K-invariant finite dimensional subspace with $V \subseteq V'$, then the restriction of $E_{V'}$ to W is 0, since otherwise there would be a non-zero element in $(W^*)^K$. Since $f - E_V(f) \in W$ for $f \in V$, we find $0 = E_{V'}(f - E_V(f)) = E_{V'}(f) - E_V(f)$. So $E_{V'}|_V = E_V$. For an arbitrary

$f \in \mathbb{C}[U]$, define $E(f) = E_V(f)$, where V is a finite dimensional K-invariant subspace containing f. For example one may take for V the linear span of the orbit through f. By the previous consideration, this is well-defined. The properties (i) and (ii) are easily checked.

Property (1) is clear by construction. Let us show (2): Choose a K-invariant finite dimensional subspace $V \subseteq \mathbb{C}[U]$ with $g \in V$. Then $V = V^K \oplus W$ for some K-invariant complement W and $E(g)$ is the projection of g onto V^K. Note that $(fW)^K = \{0\}$, since $(W^*)^K = \{0\}$. We have $fV = fV^K \oplus fW$ and $fV^K = (fV)^K$. Thus, $E(fg)$ is the projection of fg onto fV^K which is $fE(g)$. □

9.3.3. Lemma. *Let K be a reductive algebraic group and U a rational affine K-variety. Then:*

(1) *$\mathbb{C}[U]^K$ is a finitely generated $\mathbb{C}$-algebra.*
(2) *Each K-orbit $\mathcal{O}$ in U is a smooth variety which is open in its closure $\overline{\mathcal{O}}$ in U. Its boundary $\overline{\mathcal{O}}\backslash\mathcal{O}$ is a union of orbits of lower dimension.*
(3) *If I is an ideal of $\mathbb{C}[U]^K$, then $(I\mathbb{C}[U])^K = I$.*
(4) *$\mathbb{C}[U]^K$ separates closed disjoint K-invariant algebraic subsets of U.*
(5) *If $x \in U$ and $K.x$ is closed, then the isotropy group K_x is reductive.*

Let $\tau_1, \dots, \tau_m$ be generators of $\mathbb{C}[U]^K$, and set $\tau = (\tau_1, \dots, \tau_m) : U \to \mathbb{C}^m$. Then:

(6) *$\tau(U)$ is the variety of relations of $\tau_1, \dots, \tau_m$.*
(7) *If $\mathcal{O}$ is a K-orbit in U, then $\overline{\mathcal{O}}$ contains a unique closed orbit.*
(8) *τ sets up a bijection between $\tau(U)$ and closed K-orbits in U.*

Proof. To (1): Let I be the ideal in $\mathbb{C}[U]$ generated by all K-invariant homogeneous polynomials on U of positive degree. By Hilbert's basis theorem, there exist finitely many homogeneous generators $\tau_1, \dots, \tau_m$ of I. Then we claim that $\mathbb{C}[U]^K = \mathbb{C}[\tau_1, \dots, \tau_m]$. The inclusion $\mathbb{C}[\tau_1, \dots, \tau_m] \subseteq \mathbb{C}[U]^K$ is clear. We show by induction on d that every homogeneous invariant polynomial f of degree d lies in $\mathbb{C}[\tau_1, \dots, \tau_m]$. This suffices to establish also the converse inclusion. In the case $d = 0$ there is nothing to prove. So let us assume $d > 0$. Then $f \in I$, and, hence,

$$f = a_1\tau_1 + a_2\tau_2 + \cdots + a_m\tau_m,$$

where $a_1, a_2, \dots, a_m \in \mathbb{C}[U]$. Let $E : \mathbb{C}[U] \to \mathbb{C}[U]^K$ be the Reynolds operator. Apply E to both sides of the above equation, then, by 9.3.2(2), we get:

$$f = E(a_1)\tau_1 + E(a_2)\tau_2 + \cdots + E(a_m)\tau_m,$$

with $E(a_1), E(a_2), \dots, E(a_m) \in \mathbb{C}[U]^K$. We may replace each $E(a_i)$ by its homogeneous part of degree $d - \deg(\tau_i)$ and thus assume that each $E(a_i)$ is homogeneous of degree strictly less than d. By induction hypothesis, $E(a_1), E(a_2), \dots, E(a_m) \in \mathbb{C}[\tau_1, \dots, \tau_m]$, and, consequently, $f \in \mathbb{C}[\tau_1, \dots, \tau_m]$.

To (2): Let $\mathcal{O} = K.x$ with $x \in U$. Since $\mathcal{O}$ is the image of the morphism $k \mapsto k.x$, the orbit $\mathcal{O}$ contains a dense open set in its closure $\overline{\mathcal{O}}$ (this is a well-known fact from algebraic geometry, see e.g. [**Bor69**]). Now K operates transitively on $\mathcal{O}$, and it evidently leaves $\overline{\mathcal{O}}$ invariant. Since $\mathcal{O}$ contains a $\overline{\mathcal{O}}$-neighborhood of one of its points, it follows from homogeneity that $\mathcal{O}$ is open in $\overline{\mathcal{O}}$. Hence $\overline{\mathcal{O}}\backslash\mathcal{O}$ is closed and of lower dimension. Moreover, $\overline{\mathcal{O}}\backslash\mathcal{O}$ is K-invariant. So it is a union of orbits of strictly lower dimension. The smoothness of $\mathcal{O}$ follows from homogeneity.

To (3): Let I be an ideal in $\mathbb{C}[U]^K$. The inclusion $I \subseteq (I\mathbb{C}[U])^K$ is obvious. Conversely, suppose that $f \in I$, $g \in \mathbb{C}[U]$, and fg is K-invariant. Let $E : \mathbb{C}[U] \to \mathbb{C}[U]^K$ be the Reynolds operator. Consequently, by 9.3.2(2):

$$fg = E(fg) = fE(g).$$

If $f \neq 0$, then we find $g = E(g)$, and, hence, $g \in \mathbb{C}[U]^K$ and $f \cdot g \in I$, since I is an ideal in $\mathbb{C}[U]^K$. If $f = 0$, then there is nothing to prove. So (3) is shown.

To (4): We prove first: If I_1 and I_2 are two K-invariant ideals in $\mathbb{C}[U]$, then we have $(I_1+I_2)^K = I_1^K + I_2^K$. For: The inclusion $I_1^K + I_2^K \subseteq (I_1+I_2)^K$ is obvious. Suppose $f \in (I_1+I_2)^K$. Then $f = f_1 + f_2$, where $f_j \in I_j$ $(j=1,2)$. We use again the Reynolds operator $E : \mathbb{C}[U] \to \mathbb{C}[U]^K$:

$$f = E(f) = E(f_1+f_2) = E(f_1) + E(f_2) \in I_1^K + I_2^K.$$

So the assertion is proved.

Let U_1 and U_2 be two disjoint closed K-invariant algebraic subsets of U. Let I_1 and I_2 be the corresponding ideals in $\mathbb{C}[U]$. Since $1 \in \mathbb{C}[U]^K = (I_1+I_2)^K = I_1^K + I_2^K$, we have $1 = f + g$, where $f \in I_1^K$ and $g \in I_2^K$. Consequently, f is an invariant which is 0 on U_1 and 1 on U_2.

To (5): This result is due to Matsushima [**Mat60**]; a proof can also be found in [**Lun73**]. But in the restricted form we will use this result here it follows immediately from theorem 9.3.9(4) (see the remark after this theorem).

To (6): Let $Z \subseteq \mathbb{C}^m$ be the variety of relations of $\tau_1, \dots, \tau_m$. Clearly $\tau(U) \subseteq Z$. Let $z \in Z$ and let $\mathfrak{m}_z$ denote the corresponding maximal ideal in $\mathbb{C}[Z] \cong \mathbb{C}[U]^K$. By (3), $\mathfrak{m}_z\mathbb{C}[U]$ is a proper ideal of $\mathbb{C}[U]$, hence we have $\tau(u) = z$ for any u in the (non-empty) zero set of $\mathfrak{m}_z\mathbb{C}[U]$. Thus, $\tau(U) = Z$.

To (7): Let $\mathcal{O}$ be a K-orbit in U. By (2), an orbit of minimal dimension in $\overline{\mathcal{O}}$ is closed. Consequently, $\overline{\mathcal{O}}$ contains a closed orbit, and this closed orbit is unique by (4): Suppose $\mathcal{O}_1$ and $\mathcal{O}_2$ were two disjoint closed orbits in $\overline{\mathcal{O}}$, then $\tau(\mathcal{O}_1) \neq \tau(\mathcal{O}_2)$, but $\tau|_{\overline{\mathcal{O}}}$ is constant, a contradiction.

To (8): Statement (8) follows from (7). $\square$

9.3.4. Let K be a reductive algebraic group and U a rational affine K-variety. Consider a K-invariant subset $U' \subseteq U$. Let U'/K denote the set of closed K-orbits in U', and we denote by $\pi_{U,K}$ the map from U to U/K which sends $u \in U$ to the unique closed orbit in $\overline{K.u}$. The map $\pi_{U,K}$ is well-defined by lemma 9.3.3(7). We say that U' is a *K-saturated* subset of U, if $\pi_{U,K}^{-1}(\pi_{U,K}(U')) = U'$, in which case $U'/K \cong \pi_{U,K}(U')$. We give U/K the quotient structure sheaf, so $\mathbb{C}[U/K] \cong \mathbb{C}[U]^K$, and, by lemma 9.3.3, U/K is an affine variety. Since all varieties we deal with are defined over $\mathbb{C}$, it follows that U/K is also a *complex analytic space*, i.e., locally given by finitely many holomorphic equations (see [**Mum88**]). In theorem 9.3.6 we will consider the quotient holomorphic structures on U/K and U'/K.

If H is a reductive algebraic subgroup of K and P is a rational affine H-variety, then we can construct (us usual) the twisted product $K \times_H P$ which is a rational affine K-variety.

Let us state a version of Luna's slice theorem [**Lun73**]. Recall that a map ϕ between smooth complex algebraic varieties is *étale*, if the differential of ϕ is everywhere an isomorphism.

9.3.5. Theorem (Luna's slice theorem). *Let K be a reductive algebraic group and consider a finite dimensional complex representation $K \to \mathrm{GL}(W)$. Let $K.x$ be a closed orbit, $x \in W$. Choose a K_x-splitting of $W \cong T_xW$ as $T_x(K.x) + N_x$ which is possible, since K_x is reductive, by lemma 9.3.3(5), and $T_x(K.x)$ is K_x-invariant (i.e., $dl_k(x)(T_x(K.x)) \subseteq T_x(K.x)$ for $k \in K_w$). Moreover, consider the canonical equivariant map*

$$\phi : K \times_{K_x} N_x \to W \ ; \ [k,n] \mapsto k.(x+n).$$

Then there exists an affine open K-saturated subset U of W and an affine open K_x-saturated neighborhood B_x of 0 in N_x such that

$$\phi : K \times_{K_x} B_x \to U$$

and

$$\bar{\phi} : (K \times_{K_x} B_x)/K \to U/K$$

are étale, where $\bar{\phi}$ denotes the map induced by ϕ. Furthermore, ϕ and the natural map $K \times_{K_x} B_x \to B_x/K_x$ induce a K-isomorphism of $K \times_{K_x} B_x$ with the fiber product $U \times_{U/K} B_x/K_x$. In particular, the map $\phi : K \times_{K_x} B_x \to U$ is isovariant, i.e., equivariant and $K_{[k,n]} = K_{\phi([k,n])}$ for all $[k,n] \in K \times_{K_x} B_x$.

This theorem is often referred to us the *algebraic slice theorem* ([**Sch80**]). Later on we will meet the *differentiable slice theorem,* namely theorem 9.4.4. For the sake of completeness we mention also the *holomorphic slice theorem* which is an immediate consequence of theorem 9.3.5, recalling that étales algebraic morphisms between complex varieties are local analytic isomorphisms ([**Mum88**]):

9.3.6. THEOREM. *Let x, ϕ, U and B_x be as in theorem 9.3.5. Then*

(1) *K_y is conjugate to a subgroup of K_x for all $y \in U$.*

Choose a K-saturated neighborhood $\tilde{B}_x$ of 0 in B_x (classical topology) such that the canonical map $\tilde{B}_x/K_x \to \tilde{U}/K$ is a complex analytic isomorphism, where $\tilde{U} = \pi_{W,K}^{-1}(\bar{\phi}((K \times_{K_x} \tilde{B}_x)/K))$. Then $\tilde{U}$ is a K-saturated neighborhood of x and

(2) *$\phi : K \times_{K_x} \tilde{B}_x \to \tilde{U}$ is biholomorphic.*

9.3.7. Next let us derive a simple consequence of Luna's slice theorem 9.3.5. We consider again a representation $K \to \mathrm{GL}(W)$ of a reductive complex algebraic group K on a complex finite dimensional vector space W. For $w \in W$ let us define

$$D(w) := \{df(w) : f \in \mathbb{C}[W]^K\}$$

and

$$\Gamma(w) := \{\lambda \in W^* : \lambda \text{ is } K_w\text{-invariant and annihilates } \mathfrak{k}.w\},$$

where $\mathfrak{k}$ denotes the Lie algebra of K. Suppose that $\tau_1, \dots, \tau_m$ generate $\mathbb{C}[W]^K$. Then the chain rule for differentiation shows that the $d\tau_i(w)$ generate $D(w)$.

Let $f \in \mathbb{C}[W]^K$. Since f is constant on the orbit $K.w$, the differential $df(w)$ annihilates $\mathfrak{k}.w = T_e l^w.\mathfrak{k} = T_w(K.w)$ (see e.g. definition 9.5.2). Moreover, $df(w)$ is K_w-invariant: Let $k \in K_w$, then $l_k^* df(w) = d(f \circ l_k)(k.w) = df(w)$. Hence, $D(w) \subseteq \Gamma(w)$.

PROPOSITION. *Suppose that $K.w$ is closed. Then, $D(w) = \Gamma(w)$.*

PROOF. By Matsushima's theorem (see lemma 9.3.3(5) and the remark after theorem 9.3.9), the isotropy group $H = K_w$ is reductive. Since $\mathfrak{k}.w = T_w(K.w)$ is H-invariant, there exists a H-invariant subspace N of W complementary to $\mathfrak{k}.w$, and there is a unique H-stable decomposition $N = N^H \oplus N_1$, where N^H is the subspace of H-invariant vectors in N. Restriction to N clearly gives an isomorphism of $\Gamma(w)$ with $(N^*)^H \cong (N^H)^*$. We shall show that the image of $D(w) \subseteq \Gamma(w)$ is already $(N^*)^H$, hence $D(w) = \Gamma(w)$.

Consider the map $\psi : N \to W$ with $\psi(n) = w + n$. Then ψ^* maps $\mathbb{C}[W]^K$ to $\mathbb{C}[N]^H$, since for $f \in \mathbb{C}[W]^K$ and $k \in H = K_w$ we have $\psi^*(f)(k.n) = (f \circ \psi)(k.n) = f(w + k.n) = f(k.(w+n)) = f(w+n) = \psi^*(f)(n)$. Let I be the ideal in $\mathbb{C}[W]^K$ of functions vanishing at w, and let J be the ideal in $\mathbb{C}[N]^H$ of functions vanishing at 0. Then, ψ^* induces a mapping $\delta\psi : I/I^2 \to J/J^2$. We are going to show that $\delta\psi$ is an isomorphism: Consider $\phi : K \times_H N \to W$ with $[k,n] \mapsto k.(w+n)$, where $K \times_H N$ is the usual twisted product. Then, $\phi([e,0]) = w$, and ϕ is étale at $[e,0]$, by Luna's slice theorem 9.3.5. Since ϕ is equivariant, it induces a morphism $\phi/K : (K \times_H N)/K \to W/K$ which sends the point $[e,0]^\#$ corresponding to $[e,0]$ to the point $w^\#$ corresponding to w. By Luna's slice theorem 9.3.5, the map ϕ/K is étale at $[e,0]^\#$, which implies that ϕ/K induces an isomorphism of the tangent (and cotangent) spaces of $(K \times_H N)/K$ and W/K at $[e,0]^\#$ and $w^\#$. But the cotangent space of W/K at $w^\#$ is isomorphic to I/I^2, and the cotangent space of

$(K \times_H N)/K \cong N/H$ at $[e,0]^\#$ is isomorphic to J/J^2 (this is well-known from algebraic geometry, see e.g. [**Bor69**]), and, finally, the isomorphism induced by ϕ/K is easily seen to be $\delta\psi$.

$$\begin{tikzcd}
\mathbb{C}[W]^K \arrow[r, "\psi^*"] & \mathbb{C}[N]^H \\
I \arrow[r] \arrow[u] \arrow[d] & J \arrow[u] \arrow[d] \\
I/I^2 \arrow[r, "\delta\psi"] & J/J^2
\end{tikzcd}$$

Since $\delta\psi : I/I^2 \to J/J^2$ is an isomorphism, we find $\psi^*(I) + J^2 = J$, and thus $\{d(h \circ \psi)(0) : h \in \mathbb{C}[W]^K\} = \{df(0) : f \in \mathbb{C}[N]^H\}$. The latter space clearly equals $(N^*)^H$, while the former space is the image of $D(w)$ in $(N^*)^H$ via restriction to N. The proof is complete. $\square$

9.3.8. Complexification of a compact Lie group. Let G be a compact Lie group. Then G carries a unique structure of a real linear algebraic group such that any finite dimensional real representation $G \to \mathrm{GL}(V)$ is automatically a morphism of real algebraic groups ([**Che55**]). Associated to G is a reductive complex algebraic group $G^{\mathbb{C}}$, the *complexification* of G. If W is a complex representation space of G, then the representation extends uniquely from G to $G^{\mathbb{C}}$, and this property characterizes $G^{\mathbb{C}}$ ([**Che46**], [**Hoc65**], [**HM59**], [**Mos55**]). Giving $G^{\mathbb{C}}$ its classical topology, one finds that G is a maximal compact subgroup of $G^{\mathbb{C}}$. If G is a real algebraic subgroup of $\mathrm{GL}(V)$, then $G^{\mathbb{C}}$ can be taken to be the Zariski closure of G in $\mathrm{GL}(V \otimes_{\mathbb{R}} \mathbb{C})$ ([**Che46**], [**Che55**]).

If L is an algebraic group, then every compact subgroup of L (classical topology) is contained in a maximal compact subgroup, all maximal compact subgroups of L are conjugate, and L is reductive if and only if it is isomorphic to the complexification of one of its maximal compact subgroups ([**Hoc65**], [**Mos55**]).

9.3.9. The following theorem is due to Kempf and Ness [**KN79**] and Dadok and Kac [**DK85**].

THEOREM. *Let W be a complex representation space of a compact Lie group G, and let $K = G^{\mathbb{C}}$ denote the complexification of G. Suppose $(.\mid.)$ is a G-invariant Hermitian form on W with associated norm $\|.\|$. Let $w \in W$ and consider the function $F_w : K \to \mathbb{R}^+$ defined by $F_w(k) = \|kw\|^2$. Then:*

(1) *All critical points of F_w occur at minima.*
(2) *F_w has a critical point if and only if the orbit $K.w$ is closed.*

Assume that $F_w(e)$ is a minimum. Then:

(3) *$G.w = \{z \in K.w : \|z\| = \|w\|\}$.*
(4) *$K_w = (G_w)^{\mathbb{C}}$.*

PROOF. Recall that the *algebraic torus* $T := \{(t_1, \dots, t_n) : t_i \in \mathbb{C}^*\}$ is the complexification of the maximal compact subgroup G_T which consists of the elements of T whose coordinates have absolute value one. Let $\mathfrak{T}$ denote the set of maximal algebraic tori T in K such that $G \cap T$ is the maximal compact subgroup G_T of T.

For a representation of an algebraic torus T on W we may decompose W uniquely as direct sum $W = \bigoplus W_\chi$, where W_χ is the χ-eigenspace of W for the characters χ of T. Recall that a *character* $\chi : T \to \mathbb{C}^*$ is a morphism of algebraic groups sending $(t_1, \dots, t_n)$ to $\prod t_i^{m_i}$, where $m_1, \dots, m_n$ are integers. Then, our Hermitian norm $\|.\|$ on W is G_T-invariant if and only if two eigenspaces with distinct characters are perpendicular: Let $z = \sum_\chi z_\chi$ be the eigendecomposition of z, where each z_χ is non-zero. As $(t_1, \dots, t_n).z = \sum_\chi \prod_i t_i^{m_i(\chi)} z_\chi$, we have the equality

$$\|(t_1, \dots, t_n).z\|^2 = \sum_{\chi_1, \chi_2} \prod_i t_i^{m_i(\chi_1)} \overline{t_i}^{m_i(\chi_2)} (z_{\chi_1} \mid z_{\chi_2}),$$

from which the asserted equivalence can be read off. Further, using this orthogonality of the eigenspaces, we obtain:

$$F_w((t_1, \dots, t_n)) = \|(t_1, \dots, t_n).w\|^2 = \sum_\chi \left\| \prod_i t_i^{m_i(\chi)} w_\chi \right\|^2 = \sum_\chi \|w_\chi\|^2 \prod_i |t_i|^{2m_i(\chi)}.$$

Since the function F_w is G_T-invariant, it induces a function $\tilde{F}_w$ on the quotient T/G_T. Now, let us introduce coordinates on T/G_T. Define $x(t_1, \dots, t_n) = (\log|t_1|, \dots, \log|t_n|) = (x_1, \dots, x_n)$ for any $(t_1, \dots, t_n) \in T$. Hence, x defines an isomorphism $T/G_T \cong \mathbb{R}^n$, and we have

$$\tilde{F}_w((x_1, \dots, x_n)) = \sum_\chi e^{\log\|w_\chi\|^2 + 2\sum_i m_i(\chi) x_i}. \tag{9.11}$$

Let us regard K as an algebraic group defined over $\mathbb{R}$ such that G is the real locus of K. Then the tori T in $\mathfrak{T}$ are exactly the maximal tori of K which are defined over $\mathbb{R}$.

To (2): The function F_w clearly has a minimum (and therefore a critical point), if the orbit $K.w$ is closed. Conversely, suppose that the orbit $K.w$ is not closed. Then we have to show that $\inf\{F_w(k) : k \in K\}$ is not a value of F_w.

Since $K.w$ is not closed, we may find a parabolic subgroup P of K such that each maximal torus T of P contains a one-parameter algebraic subgroup $\lambda_T : \mathbb{C}^* \to T$ such that the limit $\lim_{t\to 0} \lambda_T(t).w$ exists in W and is a point outside of the orbit $K.w$. See [**Bir71**] or [**Kem78**]. Recall that a *parabolic subgroup* is a closed subgroup which contains a *Borel subgroup*, i.e., a closed, connected, solvable subgroup which is maximal for these properties. A *one-parameter algebraic subgroup* of an algebraic group H (defined over $\mathbb{C}$) is a morphism $\lambda : \mathbb{C}^* \to H$ of algebraic groups. If a morphism $f : \mathbb{C}^* \to X$ of algebraic varieties extends to a morphism $\tilde{f} : \mathbb{C} \to X$, then $y = \tilde{f}(0)$ is called the *specialization* of $f(t)$ as t specializes to 0, and we write $\lim_{t\to 0} f(t) = y$.

Let $\overline{P}$ be the parabolic subgroup of K which is conjugate to P under the real structure on K. Then $P \cap \overline{P}$ is a subgroup of K which is defined over $\mathbb{R}$, and it must contain a maximal torus S defined over $\mathbb{R}$. As a maximal torus of the intersection of two parabolic subgroups of K is a maximal torus of K (this is a consequence of the Bruhat decomposition, see [**Bor69**]), S is a maximal torus in the collection $\mathfrak{T}$.

Via the one-parameter subgroup λ_S of S which is provided by the statement above, we have an action of $\mathbb{C}^*$ on W such that the $\mathbb{C}^*$-orbit through w is not closed. Moreover, the maximal compact subgroup of $\mathbb{C}^*$, namely S^1, preserves the Hermitian form $(. \mid .)$ on W. It will suffice to prove that $\inf\{F_w(k) : k \in K\}$ is not a value of F_w when $K = \mathbb{C}^*$, since this special case implies the general case. Better yet, we may even assume that $\lim_{t\to 0} t * w$ exists in W and does not equal w, where $*$ denotes the action of $\mathbb{C}^*$ on W.

By equation (9.11), we may write the induced function $\tilde{F}_w$ on $\mathbb{C}^*/S^1 \cong \mathbb{R}$ uniquely in the form $\tilde{F}_w(x) = \sum_i a_i e^{l_i x}$ with positive a_i and increasing real l_i. Since $\lim_{t\to 0} t * w$ exists in W, the limit $\lim_{x\to -\infty} \tilde{F}_w(x)$ exists in $\mathbb{R}$, and, consequently, the l_i are non-negative. As $\lim_{t\to 0} t * w$ does not equal w, at least one l_i must be positive. Thus, $\tilde{F}_w$ is a strictly increasing function on $\mathbb{R}$. That means that $\tilde{F}_w$ and F_w never obtain a minimum value. This completes the proof of (2).

We denote by $\mathfrak{g}$ and $\mathfrak{g}_w$ the Lie algebras of G and G_w. Recall that then $\mathfrak{k} = \mathfrak{g} \oplus i\mathfrak{g}$, where $\mathfrak{k}$ denotes the Lie algebra of K. Let T_0 be a maximal torus of G and $i\mathfrak{t}$ its Lie algebra. Then we have the decomposition $K = GTG$, where T is a connected real subgroup of K with Lie algebra $\mathfrak{t}$. We denote by $T^{\mathbb{C}}$ the (maximal) complex torus in K corresponding to $\mathfrak{t}^{\mathbb{C}} = \mathfrak{t} \oplus i\mathfrak{t}$. The *weights* of the representation $K \to \mathrm{GL}(W)$ thus live in $\mathfrak{t}^*$, and we have the weight space decomposition $W = \bigoplus_{\lambda \in \mathfrak{t}^*} W_\lambda$.

Let $Y \in i\mathfrak{g}$. We shall examine the function $a(s) := \| \exp(sY).w\|^2$ for $s \in \mathbb{R}$.

CLAIM. *Either $Y.w = 0$, or $a''(s) > 0$ for all $s \in \mathbb{R}$.*

There exists a maximal toral subalgebra $\mathfrak{t} \subseteq \mathfrak{g}$ such that $Y \in i\mathfrak{t}$. Let $w = \sum_{\lambda \in \mathfrak{t}^*} w_\lambda$ be the decomposition of w in terms of weight vectors, according to the weight space decomposition $W = \bigoplus_{\lambda \in \mathfrak{t}^*} W_\lambda$ of the representation $K \to \mathrm{GL}(W)$. We have, by the orthogonality noticed above,

$$a(s) = \| \exp(sY).w\|^2 = \left\| \sum_{\lambda \in \mathfrak{t}^*} \exp(sY).w_\lambda \right\|^2 = \sum_{\lambda \in \mathfrak{t}^*} a_\lambda e^{sb_\lambda},$$

where $a_\lambda = \|w_\lambda\|^2 \geq 0$ and $b_\lambda = 2\lambda(Y) \in \mathbb{R}$. Consequently, we find

$$a''(s) = \sum_{\lambda \in \mathfrak{t}^*} a_\lambda b_\lambda^2 e^{sb_\lambda}.$$

It is thus clear that the only way in which $a''(s)$ can fail to be strictly positive is, if either $w = 0$ or if $\lambda(Y) = 0$ for all weights λ. But then $Y.w = 0$, and the claim is proved.

To (3): Let $k \in K$. Then, by the Cartan decomposition, k can be uniquely written in the form $k = g\exp(Y)$, where $g \in G$ and $Y \in i\mathfrak{g}$ (see e.g. [**Kna96**]).

The inclusion $G.w \subseteq \{z \in K.w : \|z\| = \|w\|\}$ is trivial, since $\|.\|$ is G-invariant. Suppose on the other hand that we have $\|w\| = \|k.w\| = \|g\exp(Y).w\| = \|\exp(Y).w\|$, consequently, $a(1) = \|\exp(Y).w\|^2 = \|w\|^2 = a(0)$. Since $F_w(e) = \|w\|^2 = a(0)$ is a minimum by assumption, we also have $a'(0) = 0$, and it follows that $a''(s) = 0$ for some s with $0 < s < 1$. By the claim, we find $Y.w = 0$, or equivalently $\exp(sY).w = w$, which means that $Y \in i\mathfrak{g}_w$, and hence $k.w = g\exp(Y).w = g.w$. So $\{z \in K.w : \|z\| = \|w\|\} \subseteq G.w$, too.

To (4): If $k \in K_w$, then $w = k.w = g\exp(Y).w = g.w$, using the notation and the conclusions of the previous paragraph. So $g \in G_w$ and $Y \in i\mathfrak{g}_w$. Thus, $K_w = G_w \exp(i\mathfrak{g}_w) = (G_w)^{\mathbb{C}}$, and (4) is shown.

To (1): Note first that, since $F_{h.w}(k) = \|kh.w\|^2 = F_w(kh)$, it suffices to assume that e is a critical point of F_w. Let $k \in K$ be arbitrary. We have to show that $F_w(k) \geq F_w(e)$. Again write $k = g\exp(Y)$, where $g \in G$ and $Y \in i\mathfrak{g}$. Then, $F_w(k) = \|k.w\|^2 = \|g\exp(Y).w\|^2 = \|\exp(Y).w\|^2 = a(1)$ and $F_w(e) = a(0)$. Since e is a critical point of F_w, we have $a'(0) = 0$. Let us now apply the claim: Either $Y.w = 0$ (or equivalently $\exp(sY).w = w$) implies that $F_w(k) = F_w(\exp(Y)) = F_w(e)$, or $a''(s) > 0$ for all s gives $F_w(k) = a(1) > a(0) = F_w(e)$. This shows (1). □

REMARK. Note that statement (4) of the forgoing theorem implies Matsushima's theorem [**Mat60**] which states that, if the orbit $K.w$ is closed, then the isotropy group K_w is reductive.

9.3.10. COROLLARY. *Adopt the setting of theorem 9.3.9. Let $w \in W$, and denote by $\mathfrak{g}$ and $\mathfrak{k}$ the Lie algebra of G and K, respectively. Then, the following statements are equivalent:*

(1) $(\mathfrak{g}.w \mid w) = 0$.
(2) $(\mathfrak{k}.w \mid w) = 0$.

(3) *There is an* $f \in \mathbb{C}[W]^K = \mathbb{C}[W]^G$ *such that* $df(w)(z) = (z \mid w)$ *for all* $z \in W$.

PROOF. The equivalence of (1) and (2) follows immediately from $\mathfrak{k} = \mathfrak{g} \oplus i\mathfrak{g}$.

(2) $\Rightarrow$ (3): Suppose $(\mathfrak{k}.w \mid w) = 0$. Then the function F_w has a critical point at $e \in K$, since $dF_w(e)(X) = 2(X.w \mid w)$ for $X \in \mathfrak{k}$. By theorem 9.3.9, $K.w$ is closed and $K_w = (G_w)^{\mathbb{C}}$. The linear functional $z \mapsto (z \mid w)$ vanishes on $\mathfrak{k}.w$ and is G_w-invariant (note $(g.z \mid w) = (g.z \mid g.w) = (z \mid w)$ for $g \in G_w$, since $(. \mid .)$ is G-invariant), hence K_w-invariant as well. Proposition 9.3.7 then gives (3).

(3) $\Rightarrow$ (2): If (3) holds, then $(\mathfrak{k}.w \mid w) = df(w)(\mathfrak{k}.w) = 0$, since f is K-invariant.

The equality $\mathbb{C}[W]^K = \mathbb{C}[W]^G$ follows from the fact that G is Zariski dense in $K = G^{\mathbb{C}}$. □

9.3.11. Let us now specialize to the case where $W = V \otimes_{\mathbb{R}} \mathbb{C} = V \oplus iV$ is the complexification of the representation space V of the compact Lie group G. We suppose that $\sigma_1, \ldots, \sigma_n$ are homogeneous generators of $\mathbb{R}[V]^G$. Let $K = G^{\mathbb{C}}$ be the complexification of G. The natural isomorphism $\mathbb{R}[V] \otimes_{\mathbb{R}} \mathbb{C} \cong \mathbb{C}[W]$ induces an isomorphism $\mathbb{R}[V]^G \otimes_{\mathbb{R}} \mathbb{C} \cong \mathbb{C}[W]^G = \mathbb{C}[W]^K$ and the natural extension of $\sigma : V \to \mathbb{R}^n$ to $\sigma : W \to \mathbb{C}^n$, i.e., σ_j considered as polynomials on W, is an orbit map for the representation $\rho^{\mathbb{C}} : K \to \mathrm{GL}(W)$, since G is Zariski dense in K. The quotient variety W/K, i.e., the affine variety with coordinate ring $\mathbb{C}[W]^K$, has then a real structure. The image $\sigma(W) \in \mathbb{C}^n$ is isomorphic to W/K, by lemma 9.3.3(8). Recall that $\sigma(V) = V/G \subseteq Z$, where $Z = \sigma(W) \cap \mathbb{R}^n$ is the set of real points in W/K. We assume that our G-invariant Hermitian form $(. \mid .)$ on W restricts to a G-invariant inner product $\langle . \mid . \rangle$ on V.

PROPOSITION. *Under these assumptions let* $z \in Z$ *and choose* $w \in W$ *such that* $K.w$ *is closed,* $\sigma(w) = z$ *and* $\|w\| \leq \|k.w\|$ *for all* $k \in K$*, where* $\|.\|$ *denotes the norm associated to* $(. \mid .)$. *Then we have* ($\overline{w}$ *denotes the complex conjugate of* w):

(1) $\overline{w} = g.w = g^{-1}.w$ *for some* $g \in G$.
(2) $K_{\overline{w}} = K_w = \overline{K_w} = (G_w)^{\mathbb{C}}$.
(3) *If* $z \in \sigma(V) = V/G$*, then* $w \in V$.

PROOF. We will use the function F_w which was defined in theorem 9.3.9.

To (1): The polynomials σ_j are real, so $\sigma(\overline{w}) = \overline{\sigma(w)} = \sigma(w)$, since $\sigma(w) = z \in \mathbb{R}^n$. Now $K.\overline{w} = \overline{K}.\overline{w} = \overline{K.\overline{w}}$ is closed, hence both w and $\overline{w}$ lie on closed orbits and have the same image in W/K. Thus $K.\overline{w} = K.w$, and since $\|\overline{w}\| = \|w\|$, we have $\overline{w} = g.w$ for some $g \in G$, by theorem 9.3.9(3), since $F_w(e)$ is a minimum by assumption. Write $w = v_1 + iv_2$ with $v_1, v_2 \in V$. Then, $v_1 - iv_2 = g.(v_1 + iv_2)$ implies $v_1 = g.v_1$ and $-v_2 = g.v_2$, consequently, $g^{-1}.w = \overline{w}$ and (1) is proved.

To (2): It is easy to see $G_{v_1} \cap G_{v_2} = G_w = G_{\overline{w}}$, if $w = v_1 + iv_2$. With theorem 9.3.9(4) we find $K_w = (G_w)^{\mathbb{C}} = (G_{\overline{w}})^{\mathbb{C}} = K_{\overline{w}}$ and moreover $K_w = \overline{K_w}$.

To (3): Let $v \in V$. Then, $(\mathfrak{g}.v \mid v) = \langle \mathfrak{g}.v \mid v \rangle = 0$, since $\mathfrak{g} \subseteq \mathfrak{o}(V)$. By corollary 9.3.10, this is equivalent to $(\mathfrak{k}.v \mid v) = 0$, whence the function F_v has a critical point at e. By theorem 9.3.9, it is a minimum, so $\|v\| \leq \|k.v\|$ for all $k \in K$, and $K.v$ is closed.

Now, if $\sigma(w) = z \in \sigma(V) = V/G$, there is a $v \in V$ such that $\sigma(v) = z$. Since both $K.w$ and $K.v$ are closed, we have $K.w = K.v$. Moreover, both F_w and F_v have a minimum at $e \in K$, thus we find $\|w\| = \|v\|$, and so, by theorem 9.3.9(3), $w \in G.v \subseteq V$. □

9.3.12. Let us maintain the assumptions made in 9.3.11 and let us use the same notation: w, g, etc. We introduce

$$D_{\mathbb{R}}(w) := \mathrm{span}_{\mathbb{R}}\{d\sigma_i(w) : 1 \leq i \leq n\},$$

and

$$\Gamma_{\mathbb{R}}(w) := \{\lambda \in \Gamma(w) : \lambda \circ g = \overline{\lambda}\}.$$

Recall that the complex conjugate $\overline{\lambda}$ of a linear form λ is defined by $\overline{\lambda}(x) := \overline{\lambda(\overline{x})}$. We may deduce the following consequence of proposition 9.3.11.

COROLLARY. *In the situation of 9.3.11, $D_{\mathbb{R}}(w) = \Gamma_{\mathbb{R}}(w)$.*

PROOF. $D_{\mathbb{R}}(w) \subseteq \Gamma_{\mathbb{R}}(w)$: For any $1 \le i \le n$ and any $x \in W$, we have

$$d\sigma_i(w)(g.x) = d(\sigma_i \circ l_g)(g^{-1}.w)(x) = d\sigma_i(g^{-1}.w)(x) = d\sigma_i(\overline{w})(x) = \overline{d\sigma_i(w)}(x),$$

since $\overline{w} = g^{-1}.w$, by proposition 9.3.11(1).

$\Gamma_{\mathbb{R}}(w) \subseteq D_{\mathbb{R}}(w)$: Note that complex conjugation gives a conjugate linear isomorphism between $\Gamma(w)$ and $\Gamma(\overline{w})$ (recall that $K_w = K_{\overline{w}}$, by proposition 9.3.11(2)). On the other hand, composition with the g-action gives a linear isomorphism between $\Gamma(w)$ and $\Gamma(\overline{w})$, by proposition 9.3.7. Composing these two isomorphisms, we obtain a conjugate linear isomorphism $\tau : \Gamma(w) \to \Gamma(w)$ whose fixed points are $\Gamma_{\mathbb{R}}(w)$. By proposition 9.3.7 and the remark immediately after the definition of $D(w)$ and $\Gamma(w)$, we have $\Gamma(w) = D(w) \cong D_{\mathbb{R}}(w) \otimes_{\mathbb{R}} \mathbb{C}$. It follows that the fixed points of τ in $\Gamma(w)$, namely $\Gamma_{\mathbb{R}}(w)$, lie in $D_{\mathbb{R}}(w)$. □

9.3.13. Proof of theorem 9.2.2. Finally, we have the necessary prerequisites to prove theorem 9.2.2. Let us shortly recapitulate its setting and its assertion. We consider a real finite dimensional representation $\rho : G \to \mathrm{O}(V)$ of a compact Lie group G on an Euclidean vector space V with inner product $\langle . \mid . \rangle$. Let $\sigma_1, \dots, \sigma_n$ be generators for the algebra $\mathbb{R}[V]^G$ of G-invariant polynomials on V. Let $\sigma := (\sigma_1, \dots, \sigma_n) : V \to \mathbb{R}^n$ be the orbit map, and let $\tilde{B}$ be the unique matrix valued polynomial on $V/\!\!/G$ such that $B = \tilde{B} \circ \sigma$, where $B(v) = \left(\langle d\sigma_i(v) \mid d\sigma_j(v)\rangle\right)_{ij}$ as defined in (9.10) (we denote also the dual inner product on V^* by $\langle . \mid . \rangle$). Then, we claim that $\sigma(V) = \{z \in V/\!\!/G : \tilde{B}(z) \ge 0\}$.

PROOF. Let $W = V \otimes_{\mathbb{R}} \mathbb{C} = V \oplus iV$ be the complexification of V, and consider $\sigma = (\sigma_1, \dots, \sigma_n)$ as a mapping from W to $\mathbb{C}^n$. Let $K = G^{\mathbb{C}}$ be the unique complexification of the compact Lie group G. Then K is a reductive complex algebraic group. We have $\mathbb{C}[W]^K \cong \mathbb{R}[V]^G \otimes_{\mathbb{R}} \mathbb{C}$ for the algebra of K-invariant polynomials on W which is generated by the σ_i. Let $\{. \mid .\}$ denote the K-invariant symmetric non-degenerate bilinear forms on W and W^* extending the G-invariant inner products $\langle . \mid . \rangle$ on V and V^*. We define $(x \mid y) := \{x \mid \overline{y}\}$ for all $x, y \in W$. Then, $(. \mid .)$ is a G-invariant Hermitian form on W, and let $\|.\|$ denote the associated norm.

Suppose that $z \in V/\!\!/G$. By proposition 9.3.11, we find a point $w = v_1 + iv_2$ in $W = V \oplus iV$ with the following properties: The orbit $K.w$ is closed, $\sigma(w) = z$, the isotropy group $K_w = (G_w)^{\mathbb{C}}$, $\|w\| \le \|k.w\|$ for all $k \in K$, and $\overline{w} = g.w$ for some $g \in G$. Consider the linear form $\lambda_1(x) := (x \mid w) = \{x \mid \overline{w}\}$ on W. Since $\|w\| \le \|k.w\|$ for all $k \in K$ implies $(\mathfrak{k}.w \mid w) = 0$, corollary 9.3.10 yields that λ_1 lies in $D(w)$. Moreover, consider the linear form $\lambda_2(x) := \{x \mid w\}$ on W. It is the differential at w of the K-invariant polynomial $x \mapsto \frac{1}{2}\{x \mid x\}$, and, therefore, it lies in $D(w)$, too. We put $\lambda := \frac{1}{2}(\lambda_2 - \lambda_1)$, and so we find

$$\lambda(x) = \frac{1}{2}(\lambda_2(x) - \lambda_1(x)) = \frac{1}{2}\big(\{x \mid w\} - \{x \mid \overline{w}\}\big) = \frac{1}{2}\{x \mid 2iv_2\} = \{x \mid iv_2\}.$$

By definition and proposition 9.3.7, we have $\lambda \in \Gamma(w)$, but even $\lambda \in \Gamma_{\mathbb{R}}(w)$ is true:

$$(\lambda \circ g)(x) = \{g.x \mid iv_2\} = \{x \mid g^{-1}.iv_2\} = \{x \mid \overline{iv_2}\} = \overline{\lambda}(x),$$

since $\overline{w} = g.w$. By corollary 9.3.12, the form λ has to lie in $D_{\mathbb{R}}(w)$, and we have

$$\{\lambda \mid \lambda\} = \{iv_2 \mid iv_2\} = -\{v_2 \mid v_2\} = -\langle v_2 \mid v_2\rangle \le 0. \tag{9.12}$$

Let us assume now that $z \in V/\!\!/G$, and, moreover, $\tilde{B}(z) \geq 0$. Use the point $w = v_1 + iv_2$ in W from the previous paragraph. Then we have $0 \leq \tilde{B}(z) = \tilde{B}(\sigma(w)) = B(w) = (\{d\sigma_i(w) \mid d\sigma_j(w)\})_{ij}$ which is Gram's matrix of the symmetric bilinear form $\{. \mid .\}$ on $D_{\mathbb{R}}(w) = \operatorname{span}_{\mathbb{R}}\{d\sigma_i(w)\}$. It is well-known from linear algebra that this implies that $\{. \mid .\}$ is positive semidefinite on $D_{\mathbb{R}}(w)$, and, since $\lambda \in D_{\mathbb{R}}(w)$, we have $\{\lambda \mid \lambda\} \geq 0$. Comparing this result with (9.12) implies that $v_2 = 0$, and, hence, $w = v_1 \in V$. So $z = \sigma(v_1) \in \sigma(V)$.

The converse inclusion $\sigma(V) \subseteq \{z \in V/\!\!/G : \tilde{B}(z) \geq 0\}$ is, for instance, a consequence of lemma 3.1.3 and remark 9.2.3(2). $\square$

9.3.14. We conclude the section with the following consideration: Let z, w, g, etc. be as in proposition 9.3.11. We do not assume that $\tilde{B}(z) \geq 0$. Note that $\Gamma(w)$ (equivalently, $\Gamma_{\mathbb{R}}(w)$) is invariant under the action of g if and only if $\Gamma(w)$ (equivalently, $\Gamma_{\mathbb{R}}(w)$) is invariant under complex conjugation. In particular, if G is finite, then $\Gamma(w) = (W^*)^{G_w}$ is g-invariant.

Proposition. *Let z, w, g, etc. be as above. Assume that $\Gamma(w)$ is g-invariant. Denote by $\Gamma_{\mathbb{R}}^+(w)$ (resp. $\Gamma_{\mathbb{R}}^-(w)$) the $+1$ (resp. -1) eigenspace of g acting on $\Gamma_{\mathbb{R}}(w)$. Then:*

(1) $\operatorname{rank} \tilde{B}(z) = \dim \Gamma(w)$.
(2) $\operatorname{signature} \tilde{B}(z) = \dim \Gamma_{\mathbb{R}}^+(w) - \dim \Gamma_{\mathbb{R}}^-(w)$.

Proof. If $\lambda \in \Gamma_{\mathbb{R}}^+(w)$, then we have $\overline{\lambda} = \lambda$, i.e., λ is real valued on V. Thus $(. \mid .)$ is positive definite on $\Gamma_{\mathbb{R}}^+(w)$, and similarly $(. \mid .)$ is negative definite on $\Gamma_{\mathbb{R}}^-(w)$. Hence $(. \mid .)$ is nondegenerate on $\Gamma_{\mathbb{R}}(w)$, and (1) and (2) follow. $\square$

Remark. Let $G = \mathrm{S}_n$ act on $V = \mathbb{R}^n$ by permuting the coordinates, and let $\sigma_1, \dots, \sigma_n$ be the elementary symmetric polynomials. Then, the above proposition yields the supplement in theorem 3.1.2: Let $z = (z_1, \dots, z_n) \in \mathbb{R}^n = V/\!\!/G$, and choose $w = (w_1, \dots, w_n) \in \mathbb{C}^n = W$ and $g \in G$ as in proposition 9.3.11. Recall that the w_j are the roots of the polynomial $P(x) = x^n - z_1 x^{n-1} + \dots + (-1)^n z_n$. Now we have $\Gamma(w)^* = W^{G_w} = \{y \in W : y_i = y_j \text{ if } w_i = w_j\}$. Therefore, $\dim \Gamma(w)$ is the number of distinct roots of P, which by the forgoing proposition equals $\operatorname{rank} \tilde{B}(z)$, where $\tilde{B}(z) = ((d\sigma_i(w) \mid d\sigma_j(w)))_{ij}$. The signature of $\tilde{B}(z)$ is $\dim \Gamma_{\mathbb{R}}^+(w) - \dim \Gamma_{\mathbb{R}}^-(w)$. Since $g.w = \overline{w}$, we see that $\dim \Gamma_{\mathbb{R}}^-(w)$ equals one-half the number of distinct non-real roots of P, and so the signature of $\tilde{B}(z)$ is the number of distinct real roots. In the formulation of theorem 3.1.2 we have chosen $s_1, \frac{1}{2}s_2, \dots, \frac{1}{n}s_n$ as generators of $\mathbb{R}[V]^G$, where the $s_i(x_1, \dots, x_n) = \sum_{j=1}^n x_j^i$ are the Newton polynomials, to build the Bezoutiant $((d(1/is_i)(w) \mid d(1/js_j)(w)))_{ij} = (s_{i+j-2})_{ij}$. But since the $d(1/is_i)(w)$ generate $D(w) = \Gamma(w)$ as well as the $d\sigma_i(w)$, the Bezoutiant and $\tilde{B}(z)$ have the same rank and signature.

9.4. The differentiable slice theorem

9.4.1. Orbit types. We consider a Lie group G acting smoothly from the left on a smooth manifold M. That is the smooth mapping $l : G \times M \to M : (g, x) \mapsto l(g, x) = l_g(x) = l^x(g) = g.x$ satisfies $l_g \circ l_h = l_{gh}$ and $l_e = \mathrm{id}_M$. We shall speak of a *G-manifold M*.

The closed subgroups of G can be partitioned into equivalence classes by the following relation:

$$H \sim H' \quad :\Longleftrightarrow \quad \text{There exists a } g \in G \text{ for which } H = gH'g^{-1}.$$

The equivalence class of H is denoted by (H), and it is referred to as *conjugacy class* of H. The conjugacy class of an isotropy group $G_x = \{g \in G : g.x = x\}$ is invariant under the action of G, i.e., $(G_x) = (G_{g.x})$; this is because $G_{g.x} = gG_xg^{-1}$, as one verifies directly. Therefore, we can assign to each orbit $G.x$ the conjugacy class (G_x) which we shall call the *orbit type* of the orbit through x. Two orbits are said to be of the same type, if they have the same orbit type.

If G is compact, we can define a partial ordering on the conjugacy classes simply by transferring the usual partial ordering '$\subseteq$' on the subgroups to the classes:

$$\begin{aligned}(H) \leq (H') \quad &:\iff \quad \text{There are } K \in (H) \text{ and } K' \in (H') \text{ with } K \subseteq K'.\\ &\iff \quad \text{There exists a } g \in G \text{ with } H \subseteq gH'g^{-1}.\end{aligned}$$

If G is not compact, this may not be antisymmetric. For compact G the antisymmetry of this relation is a consequence of the following lemma.

LEMMA. *Let G be a compact Lie group and H a closed subgroup of G, then $gHg^{-1} \subseteq H$ implies $gHg^{-1} = H$.*

PROOF. By iteration, $gHg^{-1} \subseteq H$ implies $g^nHg^{-n} \subseteq H$ for all $n \in \mathbb{N}$. Let us consider the set $A := \{g^n : n \in \mathbb{N}_0\}$. We shall show that g^{-1} is contained in its closure $\overline{A}$.

Suppose first that e is an accumulation point of A. Then for any neighborhood U of e there is a $n > 0$ such that $g^n \in U$. Consequently, $g^{n-1} \in g^{-1}U \cap A$. Since the sets $g^{-1}U$ form a neighborhood basis of g^{-1}, we see that g^{-1} is an accumulation point of A as well. So $g^{-1} \in \overline{A}$.

Now suppose that e is isolated in A. Then, by the compactness of G, A is finite. Therefore, $g^n = e$ for some $n > 0$, and so $g^{n-1} = g^{-1} \in A$.

Since the mapping $\mathrm{conj} : G \times G \to G : (g, h) \mapsto ghg^{-1}$ is continuous, H is closed, and $\mathrm{conj}(A, H) \subseteq H$, as we have seen at the beginning of the proof, we have $\mathrm{conj}(\overline{A}, H) \subseteq H$. In particular, $g^{-1}Hg \subseteq H$ which together with our premise implies that $gHg^{-1} = H$. □

9.4.2. DEFINITION. Let M be a G-manifold. The orbit $G.x$ is called *principal orbit*, if there is a G-invariant open neighborhood U of x in M and for all $y \in U$ a smooth equivariant map $f : G.x \to G.y$.

We call $x \in M$ a *regular point*, if $G.x$ is a principal orbit. Otherwise, x is called *singular*.

Note that the equivariant map $f : G.x \to G.y$ in this definition is automatically surjective: Let $f(x) = a.y$. For arbitrary $z = g.y \in G.y$ this gives us $z = g.y = ga^{-1}a.y = ga^{-1}.f(x) = f(ga^{-1}.x)$.

The existence of f in the above definition is equivalent to the following condition: $G_x \subseteq aG_ya^{-1}$ for some $a \in G$. For:

($\Rightarrow$): $g \in G_x$ implies $g.f(x) = f(g.x) = f(x)$. For $f(x) = a.y$ it gives $ga.y = a.y$, whence $g \in G_{a.y} = aG_ya^{-1}$.

($\Leftarrow$): Define $f : G.x \to G.y$ explicitly by $f(g.x) := ga.y$. Then we have to check that $g_1.x = g_2.x$, i.e., $g := g_2^{-1}g_1 \in G_x$, implies $g_1a.y = g_2a.y$ or $g \in G_{a.y} = aG_ya^{-1}$. This is guaranteed by our assumption. Equivariance of f follows directly from its definition.

9.4.3. Slices.

DEFINITION. Let M be a G-manifold and $x \in M$, then a subset $S \subseteq M$ is called a *slice at x*, if there is a G-invariant open neighborhood U of $G.x$ in M and a smooth equivariant retraction $r : U \to G.x$ such that $S = r^{-1}(x)$.

We can find the following properties of slices:

PROPOSITION. *If M is a G-manifold and S a slice at $x \in M$, then:*

(1) *$x \in S$ and $G_x.S \subseteq S$.*
(2) *$g.S \cap S \neq \emptyset$ implies $g \in G_x$.*
(3) *$G.S = \{g.s : g \in G, s \in S\} = U$.*
(4) *S is a G_x-manifold.*
(5) *$G_s \subseteq G_x$ for all $s \in S$.*
(6) *If $G.x$ is a principal orbit and G_x is compact, then $G_y = G_x$ for all $y \in S$ if the slice S at x is chosen small enough. In other words, all orbits near $G.x$ are principal as well.*
(7) *If two G_x-orbits $G_x.s_1$ and $G_x.s_2$ in S have the same orbit type as G_x-orbits in S, then $G.s_1$ and $G.s_2$ have the same orbit type as G-orbits in M.*
(8) *$S/G_x \cong G.S/G$ is an open neighborhood of $G.x$ in the orbit space M/G.*

Proof. Let $r : U \to G.x$ be the corresponding retraction.

To (1): $x \in S$ is clear, since $S = r^{-1}(x)$ and $r(x) = x$. To show that $G_x.S \subseteq S$, take an $s \in S$ and a $g \in G_x$. Then $r(g.s) = g.r(s) = g.x = x$, and therefore $g.s \in r^{-1}(x) = S$.

To (2): $g.S \cap S \neq \emptyset$ implies $g.s \in S$ for some $s \in S$. Then we have $x = r(g.s) = g.r(s) = g.x$, i.e., $g \in G_x$.

To (3): Since r is defined on U only and since U is G-invariant, we find $G.S = G.r^{-1}(x) \subseteq G.U = U$. For the inverse inclusion, we consider $y \in U$ with $r(y) = g.x$. We write $y = g.(g^{-1}.y)$, where $g^{-1}.y \in S$, since $r(g^{-1}.y) = g^{-1}.r(y) = g^{-1}g.x = x$. So $y \in G.S$.

To (4): This is clear from (1).

To (5): Let $g \in G_s$ for $s \in S$. Then, $g.s = s \in S$ and thus, by (2), $g \in G_x$.

To (6): By (5) we have $G_y \subseteq G_x$, so G_y is compact as well. Because $G.x$ is principal it follows that for $y \in S$ close to x, G_x is conjugate to a subgroup of G_y (see remarks after definition 9.4.2), $G_y \subseteq G_x \subseteq gG_yg^{-1}$. Since G_y is compact, $G_y \subseteq gG_yg^{-1}$ implies $G_y = gG_yg^{-1}$, by lemma 9.4.1. Therefore, $G_y = G_x$, and $G.y$ is a principal orbit, too.

To (7): For any $s \in S$ it holds that $(G_x)_s = G_s$. Namely, $(G_x)_s \subseteq G_s$ is evident; conversely, we have $G_s \subseteq G_x$, by (5), and consequently, $G_s = (G_s)_s \subseteq (G_x)_s$. So $(G_x)_{s_1} = g(G_x)_{s_2}g^{-1}$ implies $G_{s_1} = gG_{s_2}g^{-1}$, and the G-orbits have the same orbit type.

To (8): The isomorphism $S/G_x \cong G.S/G$ is given by the map $G_x.s \mapsto G.s$ (it is an injection by (2) and evidently a surjection). Since, by (3), $G.S = U$ is an open G-invariant neighborhood of $G.x$ in M, we find that $G.S/G$ is an open neighborhood of $G.x$ in M/G. □

9.4.4. The differentiable slice theorem. The following theorem (due to Koszul [**Kos53**] even though in a different version) is usually referred to as the *differentiable slice theorem*, since there also exist the *algebraic slice theorem* 9.3.5 and the *holomorphic slice theorem* 9.3.6, see [**Sch80**]. It provides a description of the G-invariant neighborhood $G.S$ of x in terms of the fiber bundle $G[S] = G \times_{G_x} S$ *associated* to the principal bundle $G \to G/G_x$:

$$\begin{tikzcd} G \times S \arrow[r, "q"] \arrow[d, "\mathrm{pr}_1"'] & G \times_{G_x} S \arrow[d, "\bar{p}"] \\ G \arrow[r, "p"] & G/G_x \end{tikzcd}$$

Then $G \times_{G_x} S$ is the orbit space $(G \times S)/G_x$ of the right action $(G \times S) \times G_x \to G \times S : ((g, s), h) \mapsto (gh, h^{-1}.s)$. The orbit of (g, s) is denoted by $[g, s]$. Note that $G \times_{G_x} S$ is a smooth G-manifold, where $h.[g, s] := [hg, s]$. Recall that q is a submersion and $(G \times S, q, G \times_{G_x} S, G_x)$ is a principal bundle.

THEOREM. *Let M be a G-manifold and S a slice at x, then there is a G-equivariant diffeomorphism of the associated bundle $G[S]$ onto $G.S$,*

$$f : G[S] = G \times_{G_x} S \to G.S$$

which maps the zero section $G \times_{G_x} \{x\}$ onto $G.x$.

PROOF. Since we have $l(gh, h^{-1}.s) = g.s = l(g, s)$ for all $h \in G_x$, there is a map $f : G[S] \to G.S$ such that the following diagram commutes:

$$\begin{tikzcd} G \times S \arrow[r, "l"] \arrow[d, "q"'] & G.S \\ G \times_{G_x} S \arrow[ur, "f"'] & \end{tikzcd}$$

f is smooth because $f \circ q = l$ is smooth and q is a submersion. It is equivariant, since l and q are equivariant:

$$f(h.[g, s]) = f(h.q(g, s)) = f(q(hg, s)) = l(hg, s) = h.l(g, s) = h.f([g, s]),$$

for $h \in G$ and $[g, s] \in G \times_{G_x} S$. Moreover, f maps the zero section $G \times_{G_x} \{x\}$ onto $G.x$. It remains to show that f is a diffeomorphism. f is bijective, since with proposition 9.4.3(2)

$$\begin{aligned} g_1.s_1 = g_2.s_2 &\Leftrightarrow s_1 = g_1^{-1}g_2.s_2 \\ &\Leftrightarrow g_1 = g_2h^{-1} \text{ and } s_1 = h.s_2 \text{ for } h = g_1^{-1}g_2 \in G_x, \end{aligned}$$

and this is equivalent to

$$q(g_1, s_1) = q(g_2, s_2).$$

The surjectivity is obvious. To see that f is a diffeomorphism let us prove that the rank of f equals the dimension of M. First of all, note that $\operatorname{rank}(l_g) = \dim(g.S) = \dim S$ and $\operatorname{rank}(l^x) = \dim(G.x)$. Since $S = r^{-1}(x)$ and $r : G.S \to G.x$ is a submersion (because $r|_{G.x} = \mathrm{id}$) it follows that $\dim(G.x) = \operatorname{codim}(S)$. Therefore,

$$\begin{aligned} \operatorname{rank}(f) = \operatorname{rank}(l) &= \operatorname{rank}(l_g) + \operatorname{rank}(l^x) \\ &= \dim S + \dim(G.x) = \dim S + \operatorname{codim} S = \dim M. \end{aligned}$$

This completes the proof. □

9.4.5. The existence of slices. After having defined slices and discussed their properties, let us investigate under which conditions they exist. As we will see at the beginning of the next section, in our setting, where G is compact and acts orthogonally on V, the existence of slices at each point $v \in V$ is quite natural. Hence, we shall present the following result concerning more general situations without proof.

THEOREM ([**Mic97**], [**Pal61**]). *Let M be a G-manifold and $x \in M$ a point with compact isotropy group G_x. If for all open neighborhoods U of G_x in G there is a neighborhood W of x in M such that $\{g \in G : g.W \cap W \neq \emptyset\} \subseteq U$, then there exists a slice at x.*

Note that the conditions of this theorem are satisfied for all $x \in M$, if M is a proper G-manifold, in particular if G is compact; see e.g. [**Mic97**].

9.4.6. Proper actions.

Definition (Definition and Lemma). A smooth action $l : G \times M \to M$ is called *proper*, if it satisfies one of the following equivalent conditions:

(1) $(l, \mathrm{pr}_2) : G \times M \to M \times M,\ (g,x) \mapsto (g.x, x)$, is a proper mapping.
(2) $g_n.x_n \to y$ and $x_n \to x$ in M, for some $g_n \in G$ and $x_n, x, y \in M$, implies that these g_n have a convergent subsequence in G.
(3) If K and L are compact in M, then $\{g \in G : g.K \cap L \neq \emptyset\}$ is compact as well.

Proof. (1) $\Rightarrow$ (2) is a direct consequence of the definitions.

(2) $\Rightarrow$ (3): Let $(g_n)_n$ be a sequence in $\{g \in G : g.K \cap L \neq \emptyset\}$ and $x_n \in K$ such that $g_n.x_n \in L$. Since K and L are compact, we can choose convergent subsequences $(x_{n_k})_k$ and $(g_{n_k}.x_{n_k})_k$. Now (2) guarantees that we can find a subsequence of $(g_{n_k})_k$, and hence of $(g_n)_n$, which is convergent in $\{g \in G : g.K \cap L \neq \emptyset\}$. Therefore, $\{g \in G : g.K \cap L \neq \emptyset\}$ is compact.

(3) $\Rightarrow$ (1): Let R be a compact subset of $M \times M$. Then $L := \mathrm{pr}_1(R)$ and $K := \mathrm{pr}_2(R)$ are compact, and $(l, \mathrm{pr}_2)^{-1}(R) \subseteq \{g \in G : g.K \cap L \neq \emptyset\} \times K$. By (3), $\{g \in G : g.K \cap L \neq \emptyset\}$ is compact. Consequently, $(l, \mathrm{pr}_2)^{-1}(R)$ is compact, and (l, pr_2) is proper. □

It is a direct consequence of (2) that for compact G every G-action is proper. Furthermore, if G acts properly on some manifold, then all isotropy groups are compact: Set $K = L = \{x\}$ in (3).

9.4.7. Theorem. *If M is a proper G-manifold, then M/G is completely regular.*

Proof. Choose $F \subseteq M/G$ closed and $\pi(x_0) \notin F$, where $\pi : M \to M/G$. Let U be a compact neighborhood of x_0 in M fulfilling $U \cap \pi^{-1}(F) = \emptyset$, and let $f \in C^\infty(M, [0,\infty))$ with support in U such that $f(x_0) > 0$. By 9.4.6(3), $\{g \in G : g.x \in \operatorname{supp} f\}$ is compact, for arbitrary $x \in M$. Hence the map $g \mapsto f(g.x)$ has compact support, and so $\tilde f : x \mapsto \int_G f(g.x) d\mu_r(g)$ is well defined, where $d\mu_r$ stands for the right Haar measure. To see that $\tilde f$ is smooth, let x_1 be a point in M and V a compact neighborhood of x_1. Then, by 9.4.6(3), the set $\{g \in G : g.V \cap \operatorname{supp} f \neq \emptyset\}$ is compact. Therefore, $\tilde f$ restricted to V is smooth, and in particular $\tilde f$ is smooth in x_1. Moreover, $\tilde f$ is G-invariant and $\tilde f(x_0) > 0$, by definition. We have $\operatorname{supp} \tilde f \subseteq G.\operatorname{supp} f \subseteq G.U$, and, consequently, $\operatorname{supp} \tilde f \cap \pi^{-1}(F) = \emptyset$. Since $\tilde f \in C^\infty(M, [0,\infty))^G$, it factors over π to a map $\bar f \in C^0(M/G, [0,\infty))$, with $\bar f(\pi(x_0)) > 0$ and $\bar f|_F = 0$. □

9.4.8. Orbits are closed submanifolds. Finally, we want to show that the orbits of a proper action are closed submanifolds. For that we need the following lemma:

Lemma. *A continuous proper map $f : X \to Y$ between two topological manifolds is closed.*

Proof. Consider a closed set $A \subseteq X$, and take a point y in the closure of $f(A)$. Let $f(a_n) \in f(A)$ converge to y. Let U be a compact neighborhood of y in Y. For all large enough n, we have $f(a_n) \in U$ and thus $a_n \in f^{-1}(U)$. Since f is proper, $f^{-1}(U)$ is compact. Hence the sequence $(a_n)_n$ has a convergent subsequence $(a_{n_k})_k$ with $a_{n_k} \to a$. We find $a \in A$, since A is closed. By continuity of f, we obtain $y = f(a) \in f(A)$. □

Proposition. *The orbits of a proper action $l : G \times M \to M$ are closed submanifolds.*

PROOF. By the preceding lemma, (l, pr_2) is closed. Therefore, $(l, \mathrm{pr}_2)(G, x) = G.x \times \{x\}$ and with it $G.x$ is closed.

Next let us show that $l^x : G \to G.x$ is an open mapping. Since l^x is G-equivariant, we only have to show that, for an open neighborhood U of e in G, $l^x(U) = U.x$ is an open neighborhood of x in $G.x$. Let as assume the contrary: There exits a sequence $(g_n.x)_n \subseteq G.x \backslash U.x$ which converges to x. Then by definition 9.4.6(2), $(g_n)_n$ has a convergent subsequence with limit $g \in G_x$. On the other hand, since $g_n.x \notin U.x = UG_x.x$, we have $g_n \notin UG_x$, and, since $UG_x = \bigcup_{g \in G_x} Ug$ is open, we have $g \notin UG_x$ as well. This contradicts $g \in G_x$, by the choice of U.

Now consider the following commuting diagram:

$$\begin{tikzcd} G \arrow[r, "l^x"] \arrow[d, "p"'] & G.x \\ G/G_x \arrow[ur, "i"'] & \end{tikzcd}$$

As the integral manifold of fundamental vector fields, $G.x$ is an initial submanifold, and i is an injective G-equivariant immersion, see e.g. [**KMS93**]. Since $i \circ p = l^x$ is open, i is open as well. Therefore, it is a homeomorphism, and $G.x$ is an embedded submanifold of M. □

9.4.9. THEOREM. *If M is a proper G-manifold, then there is a G-invariant Riemannian metric on M.*

PROOF. Since the manifold M has countable base, there exists an exhaustion $(K_j)_{j \in \mathbb{N}}$ by compact sets K_j of M, i.e., $K_j \subseteq \operatorname{int} K_{j+1}$ for all $j \in \mathbb{N}$, where $\operatorname{int} K$ denotes the topological interior of K, and $M = \bigcup_{j \in \mathbb{N}} K_j$. For any $j \in \mathbb{N}$ choose a function $f_j \in C^\infty(M, [0,1])$ such $\operatorname{supp} f_j \subseteq \operatorname{int} K_{j+1}$ and $f_j|_{K_j} = 1$.

Let γ be an arbitrary Riemannian metric on M. By means of the right Haar measure $d\mu_r$ we define G-invariant smooth sections $\gamma_j : M \to T^*M \otimes_x T^*M$ by

$$\gamma_j(x)(X_x, Y_x) := \int_G f_j(g.x) \gamma_{g.x}(T_x l_g.X_x, T_x l_g.Y_x) d\mu_r(g),$$

for $x \in M$ and $X_x, Y_x \in T_x M$. Note that the map $g \mapsto f_j(g.x)$ has compact support. By the assumptions on γ and f_j all forms $\gamma_j(x)$ are symmetric and positive semidefinite. If $x \in G.K_j$, then $\gamma_j(x)$ is even positive definite. The family $(U_j)_{j \in \mathbb{N}}$ with $U_j = G. \operatorname{int} K_j$ describes a G-invariant open covering of M. There exists a G-invariant partition of unity $(\phi_j)_{j \in \mathbb{N}}$ subordinate to $(U_j)_{j \in \mathbb{N}}$, by the lemma below. Then

$$\tilde{\gamma} := \sum_{j \in \mathbb{N}} \phi_j \gamma_j$$

gives a G-invariant Riemannian metric on M. □

LEMMA. *Let M be a proper G-manifold. To any covering of M by G-invariant open sets there exists a subordinate partition of unity by G-invariant smooth functions.*

PROOF. Assume that $(U_j)_{j \in J}$ is a covering of M by G-invariant open sets. Since M/G is paracompact, there exists a locally finite covering of M/G by open sets V_j such that $\pi^{-1}(V_j) \subseteq U_j$, where $\pi : M \to M/G$. There exists a locally finite partition of unity $(\psi_i)_{i \in \mathbb{N}}$ on M and a mapping $\iota : \mathbb{N} \to J$ such that $\operatorname{supp} \psi_i$ is compact for all $i \in \mathbb{N}$ and such that $\operatorname{supp} \psi_i \subseteq \pi^{-1}(V_{\iota(i)})$. Since $\operatorname{supp} \psi_i$ is compact, for $x \in M$ and $i \in \mathbb{N}$, the following integral exists:

$$\psi_i^G(x) := \int_G \psi_i(g.x) d\mu_r(g),$$

where $d\mu_r$ denotes the right Haar measure. Then ψ_i^G is a smooth function on M and $\operatorname{supp}\psi_i^G \subseteq \pi^{-1}(V_{\iota(i)})$.

The family of supports $(\operatorname{supp}\psi_i^G)_{i\in\mathbb{N}}$ need not be locally finite anymore. Since $C^\infty(M)$ is a Fréchet space, there exists a sequence of seminorms $\|.\|_i$ on $C^\infty(M)$ defining the Fréchet topology such that $\|.\|_i \leq \|.\|_{i+1}$ for all $i \in \mathbb{N}$. Let us define

$$\tilde{\phi}_j := \sum_{\substack{i\in\mathbb{N}\\ \iota(i)=j}} \frac{1}{2^i\|\psi_i^G\|_i}\psi_i^G.$$

Then the functions $\tilde{\phi}_j$ are smooth, G-invariant, and satisfy $\operatorname{supp}\tilde{\phi}_j \subseteq \pi^{-1}(V_j)$. Since the covering $(V_j)_{j\in J}$ is locally finite, the family of supports $\operatorname{supp}\tilde{\phi}_j$ is locally finite as well.

Therefore, we may define, for all $x \in M$,

$$\phi_j(x) = \frac{1}{\tilde{\phi}(x)}\tilde{\phi}_j(x) \quad \text{where } \tilde{\phi}(x) = \sum_{j\in J}\tilde{\phi}_j(x).$$

Then $(\phi_j)_{j\in J}$ is a locally finite G-invariant partition of unity subordinate to $(U_j)_{j\in J}$. □

9.5. Reducing the problem

9.5.1. Normal slices.

Definition. Let M be a proper *Riemannian* G-manifold, i.e., endowed with a G-invariant Riemannian metric. Let $x \in M$. The *normal bundle to the orbit* $G.x$ is defined as

$$\operatorname{Nor}(G.x) := T(G.x)^\perp.$$

Let $\operatorname{Nor}_\epsilon(G.x) := \{X \in \operatorname{Nor}(G.x) : |X| < \epsilon\}$, and choose $r > 0$ small enough for $\exp_x : T_xM \supseteq B_r(0_x) \to M$ to be a diffeomorphism onto its image and for $\exp_x(B_r(0_x)) \cap G.x$ to have only one component, where $B_r(0_x)$ is the open ball with radius r centered at $0_x \in T_xM$. Then, since the action of G is isometric, exp defines a diffeomorphism from $\operatorname{Nor}_{\frac{r}{2}}(G.x)$ onto an open neighborhood of $G.x$, so $\exp(\operatorname{Nor}_{\frac{r}{2}}(G.x)) =: U_{\frac{r}{2}}$ is a tubular neighborhood of $G.x$. We define the *normal slice at* x by

$$S_x := \exp_x(\operatorname{Nor}_{\frac{r}{2}}(G.x))_x.$$

Proposition. *The so defined normal slice $S_x = \exp_x(\operatorname{Nor}_{\frac{r}{2}}(G.x))_x$ at x is indeed a slice at x and satisfies $S_{g.x} = g.S_x$.*

Proof. Let us check first that S_x satisfies the mentioned equation:

$$\begin{aligned} S_{g.x} = \exp_{g.x}(\operatorname{Nor}_{\frac{r}{2}}(G.x))_{g.x} &= \exp_{g.x}(T_xl_g(\operatorname{Nor}_{\frac{r}{2}}(G.x))_x) \\ &= l_g(\exp_x(\operatorname{Nor}_{\frac{r}{2}}(G.x))_x) = g.S_x, \end{aligned}$$

since G acts isometrically. Recall here that for isometries ϕ we have $\phi(\exp_x(tX)) = \exp_{\phi(x)}(tT_x\phi.X)$ which is due to the fact that isometries map geodesics to geodesics, and the starting vector of the geodesic $t \mapsto \phi(\exp_x(tX))$ is $T_x\phi.X$.

Consider the mapping $r : G.S_x = \bigcup_{g\in G} S_{g.x} \to G.x : \exp_{g.x} X \mapsto g.x$. It is smooth, equivariant,

$$r(l_h(\exp_{g.x} X)) = r(\exp_{hg.x}(T_xl_h.X)) = hg.x = l_h(r(\exp_{g.x} X)),$$

and a retraction

$$r(r(\exp_{g.x} X)) = r(g.x) = r(\exp_{g.x} 0_{g.x}) = g.x = r(\exp_{g.x} X).$$

Moreover, $r^{-1}(x) = S_x$, making it a slice at x. □

9.5.2. Definition. Let M be a G-manifold and $x \in M$, then there is a representation of the isotropy group G_x

$$G_x \to \mathrm{GL}(T_xM) : g \mapsto T_x l_g$$

called the *isotropy representation.*

If M is a Riemannian G-manifold, then the isotropy representation is orthogonal, and $T_x(G.x)$ is an invariant subspace under G_x. So $N_x := T_x(G.x)^\perp = \mathrm{Nor}(G.x)_x$ is also G_x-invariant, and

$$G_x \to \mathrm{O}(N_x) : g \mapsto T_x l_g$$

is called the *slice representation.*

Observe that $T_x(G.x) = T_e l^x.\mathfrak{g}$ where $\mathfrak{g} = \mathrm{Lie}(G)$, the Lie algebra of G. For: $X \in T_x(G.x) \Leftrightarrow X = \frac{d}{dt}\big|_{t=0} c(t)$ for some smooth curve $c(t) = g_t.x \in G.x$ with $g_0 = e$, i.e., $X = \frac{d}{dt}\big|_{t=0} l^x(g_t) \in T_e l^x.\mathfrak{g}$.

9.5.3. Reducing the problem to the slice representation. We adopt the setting in 8.2, where a compact Lie group G acts orthogonally on a real finite dimensional vector space V. Fix a point $v \in V$, and consider the normal slice S_v which is an open ball centered at 0 in the normal subspace $N_v = T_v(G.v)^\perp$ of the orbit $G.v$ through v. Then we recall proposition 9.4.3(8) and theorem 9.4.4, where now we can replace 'smooth' by 'real analytic' anywhere, since the vector space V is a real analytic manifold and the G-action on V is real analytic, too. Consequently, there exists a G-invariant neighborhood U of v in V which is real analytically G-isomorphic to the associated bundle $G \times_{G_v} S_v$, and the quotient U/G is homeomorphic to S_v/G_v.

In view of the lifting problem described in section 8.1 it follows:

Theorem. *The problem of local lifting curves in V/G passing through $\sigma(v)$ reduces to the same problem for curves in N_v/G_v passing through* 0.

9.5.4. The set of regular points lies open and dense. Recall the definition of a regular point given in definition 9.4.2. We give now other characterizations in terms of slices and slice representations:

Lemma. *Let M be a Riemannian G-manifold, where G is a compact Lie group, and let $x \in M$. Then the following statements are equivalent:*

(1) *x is a regular point.*
(2) *The slice representation at x is trivial.*
(3) *$G_y = G_x$ for all $y \in S_x$ for a sufficiently small slice S_x at x.*

Proof. Clearly, (2) $\Leftrightarrow$ (3). To see (3) $\Rightarrow$ (1) let S_x be a small slice at x such that (3) holds. Then $U := G.S_x$ is an invariant open neighborhood of $G.x$ in M, and for all $g.s \in U$ we have $G_{g.s} = gG_sg^{-1} = gG_xg^{-1}$. Therefore, $G.x$ is a principal orbit; see remarks after definition 9.4.2. The converse is true by proposition 9.4.3(6), since G_x is compact. □

Let us return to our setting. Assume $v \in V$ is regular. By theorem 9.4.4 and the previous lemma, there is a neighborhood of v which is analytically G-isomorphic to $G/G_v \times S_v \cong G.v \times S_v$. *The set V_{reg} of regular points in V is open and dense in V:*

- Suppose $v \in V_{\mathrm{reg}}$. There is a slice S_v at v, and, by proposition 9.4.3(6), S_v can be chosen small enough for all orbits through S_v to be principal as well. Hence $G.S_v$ is an open neighborhood of v in V_{reg} (by proposition 9.4.3(3)).

- To see that V_{reg} is dense in V, let $U \subseteq V$ be open, $x \in U$, and S_x a slice at x. We shall show that then U contains a regular point. Choose a point $y \in G.S_x \cap U$ for which the isotropy group G_y has minimal dimension and smallest number of connected components for this dimension in all of $G.S_x \cap U$ (remember G and hence all isotropy groups are compact). Let S_y be a slice at y. Then $G.S_x \cap G.S_y \cap U$ is open, and for any $z \in G.S_x \cap G.S_y \cap U$ we have $z \in g.S_y = S_{g.y}$ (by proposition 9.5.1) for a $g \in G$. Consequently, $G_z \subseteq G_{g.y} = gG_yg^{-1}$, by proposition 9.4.3(5). By choice of y, this implies $G_z = gG_yg^{-1}$ for all $z \in G.S_x \cap G.S_y \cap U$, and $G.y$ is a principal orbit.

Remark. More generally, for a proper G-manifold M the set of regular points M_{reg} is open and dense in M, too. The proof is essentially the same as in the special case above.

9.5.5. The space $N_v^{G_v}$. Finally, we can give a description of the subspace $N_v^{G_v}$ of G_v-invariant vectors in N_v in terms of the generators $\sigma_1, \dots, \sigma_n$ of $\mathbb{R}[V]^G$. It is due to Sartori [**Sar83**].

Theorem. *Consider a real finite dimensional representation $\rho : G \to \mathrm{O}(V)$ of a compact Lie group G on an Euclidean vector space V with inner product $\langle . \mid . \rangle$. Let $\sigma_1, \dots, \sigma_n$ be generators for the algebra $\mathbb{R}[V]^G$ of G-invariant polynomials on V. Let $\sigma := (\sigma_1, \dots, \sigma_n) : V \to \mathbb{R}^n$ be the orbit map. Then, for $v \in V$, the subspace $N_v^{G_v}$ of G_v-invariant vectors in $N_v = T_v(G.v)^\perp$ is spanned as a real vector space by the gradients* $\operatorname{grad}\sigma_1(v), \dots, \operatorname{grad}\sigma_n(v)$.

Proof. Clearly each $\operatorname{grad}\sigma_i(v) \in N_v^{G_v}$. Namely,

$$\langle T_v(G.v) \mid \operatorname{grad}\sigma_i(v)\rangle = d\sigma_i(v)(T_v(G.v)) = 0,$$

since σ_i is constant on $G.v$, and, for arbitrary $w \in V$ and $g \in G_v$,

$$\begin{aligned}\langle T_v l_g.w \mid T_v l_g.\operatorname{grad}\sigma_i(v)\rangle &= \langle w \mid \operatorname{grad}\sigma_i(v)\rangle = d\sigma_i(v)(w) \\ &= d(\sigma_i \circ l_g)(v)(w) = d\sigma_i(v)(T_v l_g.w) = \langle T_v l_g.w \mid \operatorname{grad}\sigma_i(v)\rangle.\end{aligned}$$

In the following we will identify G with its image $\rho(G) \subseteq \mathrm{O}(V)$. Its Lie algebra $\mathfrak{g}$ is then a subalgebra of $\mathfrak{o}(V)$ and can be realized as a Lie algebra consisting of skew-symmetric matrices.

Let $v \in V$, and let S_v be the normal slice at v which is chosen so small that the projection of the tubular neighborhood $p_{G.v} : G.S_v \to G.v$ (see theorem 9.4.4) from the diagram

$$\begin{tikzcd} G \times S_v \arrow[r, "q"] & G \times_{G_v} S_v \arrow[r, "\pi"] \arrow[d, "p"'] & G.S_v \arrow[d, "p_{G.v}"] \\ & G/G_v \arrow[r, "\pi", "\cong"'] & G.v \end{tikzcd}$$

has the property, that for any $w \in G.S_v$ the point $p_{G.v}(w) \in G.S_v$ is the unique point in the orbit $G.v$ which minimizes the distance between w and the orbit $G.v$. Remember that each orbit is closed.

Choose $n \in N_v^{G_v}$ so small that $x := v + n \in S_v$. Hence $p_{G.v}(x) = v$. For the related isotropy groups we find $G_x \subseteq G_v$, by proposition 9.4.3(5). On the other hand we have $G_v \subseteq G_v \cap G_n \subseteq G_x$, so $G_x = G_v$. Let S_x be the normal slice at x, chosen so small that $p_{G.x} : G.S_x \to G.x$ has the same minimizing property as $p_{G.v}$ above, but so large that $v \in G.S_x$ (choose n smaller if necessary). Then we find

$p_{G.x}(v) = x$, since for the Euclidean distance in V we have

$$\begin{aligned} |v - x| &= \min_{g\in G} |g.v - x| && \text{since } v = p_{G.v}(x) \\ &= \min_{g\in G} |v - g^{-1}.x| && \text{since } G \text{ acts orthogonally.} \end{aligned}$$

For $w \in G.S_x$ we consider the local, smooth, G-invariant function given by

$$\begin{aligned} w \mapsto \operatorname{dist}(w, G.x)^2 &= \operatorname{dist}(w, p_{G.x}(w))^2 \\ &= \langle w - p_{G.x}(w) \mid w - p_{G.x}(w)\rangle \\ &= \langle w \mid w\rangle + \langle p_{G.x}(w) \mid p_{G.x}(w)\rangle - 2\langle w \mid p_{G.x}(w)\rangle \\ &= \langle w \mid w\rangle + \langle x \mid x\rangle - 2\langle w \mid p_{G.x}(w)\rangle. \end{aligned}$$

Its derivative is

$$d(\operatorname{dist}(\quad, G.x)^2)(w)(y) = 2\langle w \mid y\rangle - 2\langle y \mid p_{G.x}(w)\rangle - 2\langle w \mid dp_{G.x}(w)(y)\rangle. \tag{9.13}$$

We shall show below that

$$\langle v \mid dp_{G.x}(v)(y)\rangle = 0 \qquad \text{for all } y \in V, \tag{9.14}$$

such that the derivative at v is given by

$$d(\operatorname{dist}(\quad, G.x)^2)(v)(y) = 2\langle v \mid y\rangle - 2\langle y \mid p_{G.x}(v)\rangle = 2\langle v - x \mid y\rangle = -2\langle n \mid y\rangle. \tag{9.15}$$

Now let us choose a smooth G_x-invariant function $f : S_x \to \mathbb{R}$ with compact support which equals 1 in an open ball around x and extend it smoothly (see the diagram above, but for S_x) to $G.S_x$ and then to the whole of V. We assume that f is still equal to 1 in a neighborhood of v. Then $g = f \cdot \operatorname{dist}(\quad, G.x)^2$ is a smooth G-invariant function on V which coincides with $\operatorname{dist}(\quad, G.x)^2$ near v. By the theorem of Schwarz 8.2.5, there is a smooth function $h \in C^\infty(\mathbb{R}^n)$ such that $g = h \circ \sigma$, where $\sigma = (\sigma_1, \dots, \sigma_n) : V \to \mathbb{R}^n$ is the orbit map. Consequently, we have, by equation (9.15),

$$\begin{aligned} -2n &= \operatorname{grad}(\operatorname{dist}(\quad, G.x)^2)(v) \\ &= \operatorname{grad} g(v) = \operatorname{grad}(h \circ \sigma)(v) \\ &= \sum_{i=1}^n \frac{\partial h}{\partial y_i}(\sigma(v)) \operatorname{grad} \sigma_i(v), \end{aligned}$$

which proves the result.

It remains to check equation (9.14). We have $T_v V = T_v(G.v) \oplus N_v$, and thus the normal space $N_x = \operatorname{Nor}(G.x)_x = \ker dp_{G.x}(v)$ is still transversal to $T_v(G.v) = T_e l^v.\mathfrak{g}$, if n is small enough. That means that it suffices to show that $\langle v \mid dp_{G.x}(v)(X.v)\rangle = 0$ for each $X \in \mathfrak{g} \subseteq \mathfrak{o}(V)$. Now $x = p_{G.x}(v)$ implies $|v - x|^2 = \min_{g\in G} |v - g.x|^2$, and therefore the derivative of $g \mapsto \langle v - g.x \mid v - g.x\rangle$ at e vanishes. Consequently, we have for all $X \in \mathfrak{g}$

$$0 = 2\langle -X.x \mid v - x\rangle = 2\langle X.x \mid x\rangle - 2\langle X.x \mid v\rangle = 0 - 2\langle X.x \mid v\rangle, \tag{9.16}$$

since the action of X on V is skew-symmetric. Let us consider the equation $p_{G.x}(g.v) = g.p_{G.x}(v)$ and differentiate it with respect to g at $e \in G$ in the direction $X \in \mathfrak{g}$ to obtain in turn

$$dp_{G.x}(v)(X.v) = X.p_{G.x}(v) = X.x,$$

and hence by (9.16)

$$\langle v \mid dp_{G.x}(v)(X.v)\rangle = \langle v \mid X.x\rangle = 0.$$

This completes the proof. $\square$

9.5.6. There is precisely one principal orbit type.

THEOREM. *If M is a proper G-manifold then the set M_{sing}/G of all singular G-orbits does not locally disconnect the orbit space M/G (that is for every point in M/G the connected neighborhoods remain connected even after removal of all singular orbits).*

PROOF. We shall reduce the statement of the theorem to an assertion about the slice representation. There is a G-invariant Riemannian metric on M, by theorem 9.4.9, which makes M a Riemannian G-manifold. Let S_x be the normal slice at x. Then S_x is again a Riemannian manifold, and the compact group G_x acts isometrically on S_x. A principal G_x-orbit in S_x is the restriction of a principal G-orbit in M: Recall that $G_x.s$ ($s \in S_x$) is principal means that all orbits in a sufficiently small neighborhood of $G_x.s$ in S_x have the same orbit type as $G_x.s$ (see lemma 9.5.4). Therefore, by proposition 9.4.3(7), the corresponding G-orbits in $G.S_x$ are also of the same type, and so $G.s$ is principal as well. It follows that there are 'fewer' singular G-orbits in $G.S_x$ than there are singular G_x-orbits in S_x. Now cover M with tubular neighborhoods like $G.S_x$, and recall that $G.S_x/G \cong S_x/G_x$ (proposition 9.4.3(8)). This together with the above argument shows that it will suffice to prove the statement of the theorem for the slice action. Consequently, we have reduced the assertion to the following:

If V is a real n-dimensional Euclidean vector space and G is a compact Lie group acting isometrically on V, then the set V_{sing}/G of all singular G-orbits in V does not locally disconnect the orbit space V/G.

We will show this by induction on the dimension n of V. For $n = 1$, that is $V = \mathbb{R}$, the only nontrivial choice for G is $\mathrm{O}(1) \cong \mathbb{Z}_2$. In this case, $\mathbb{R}/G \cong [0, \infty)$ and $\mathbb{R}_{\text{sing}}/G = \{0\}$. Clearly, $\{0\}$ does not locally disconnect $[0, \infty)$.

Now suppose the assertion is true for all dimensions smaller than n. Since the action of G on V is isometric, the submanifold S^{n-1} is left invariant and we may consider the induced G-action on S^{n-1}. For any $x \in S^{n-1}$ we can apply the induction hypothesis to the corresponding slice representation $G_x \to \mathrm{O}(\mathrm{Nor}(G.x)_x)$, since the dimension of $\mathrm{Nor}(G.x)_x$ is evidently smaller than n. This implies for the G_x-action on $S_x := \exp_x(\mathrm{Nor}_{\frac{r}{2}}(G.x))_x$ (for $r > 0$ sufficiently small) that $S_x/G_x \cong G.S_x/G$ is not locally disconnected by its singular points. We may cover S^{n-1} by tubular neighborhoods like $G.S_x$, and hence we see that all of S^{n-1}/G is not locally disconnected by its singular orbits. Since homotheties on V commute with the G-action on V and thus they are G-equivariant diffeomorphisms, the singular orbits in V (not including $\{0\}$) project radially onto singular orbits in S^{n-1}. So if we view the unit ball $D^n := \{x \in V : |x| \leq 1\}$ as cone over S^{n-1} and denote the cone construction by

$$\mathrm{cone}(S^{n-1}) := \{[x, t] : x \in S^{n-1} \text{ and } 0 \leq t \leq 1\}/\{[x, 0] : x \in S^{n-1}\},$$

then we have $D^n_{\text{sing}} = \mathrm{cone}(S^{n-1}_{\text{sing}})$. Moreover, we have a homeomorphism

$$\mathrm{cone}(S^{n-1})/G \longrightarrow \mathrm{cone}(S^{n-1}/G) : G.[x, t] \longmapsto [G.x, t],$$

since the G-action preserves the 'radius' t. Therefore,

$$D^n/G = \mathrm{cone}(S^{n-1})/G \cong \mathrm{cone}(S^{n-1}/G)$$

and

$$D^n_{\text{sing}}/G = \mathrm{cone}(S^{n-1}_{\text{sing}})/G \cong \mathrm{cone}(S^{n-1}_{\text{sing}}/G).$$

Since S^{n-1}_{sing}/G does not locally disconnect S^{n-1}/G as seen before, we also find that $D^n_{\text{sing}}/G \cong \mathrm{cone}(S^{n-1}_{\text{sing}}/G)$ does not locally disconnect $D^n/G \cong \mathrm{cone}(S^{n-1}/G)$. This completes the induction and hence the proof. □

Corollary. *Let M be a connected proper G-manifold. Then the orbit space M/G is connected, too, and M has precisely one principal orbit type.*

Proof. Since M is connected and the quotient map $\pi : M \to M/G$ is continuous, its image M/G is connected as well. By theorem 9.5.6, we have that $(M/G)\backslash(M_{\text{sing}}/G) = M_{\text{reg}}/G$ is connected. On the other hand, by lemma 9.5.4, the orbits of a certain principal orbit type form an open subset of M/G, in particular of M_{reg}/G. Therefore, if there were more than one principal orbit type, these orbit types would partition M_{reg}/G into disjoint nonempty open subsets, contradicting the fact that M_{reg}/G is connected. □

9.5.7. Corollary. *Let M be a connected proper G-manifold of dimension n and let k be the least number of connected components of all isotropy groups of dimension $m := \inf\{\dim G_x | x \in M\}$. Then, for $x_0 \in M$, the following two conditions are equivalent:*

(1) *$G.x_0$ is a principal orbit.*
(2) *The isotropy group G_{x_0} has dimension m and k connected components, i.e., G_{x_0} is smallest possible.*

If moreover G is connected and simply connected, these conditions are again equivalent to:

(3) *The orbit $G.x_0$ has dimension $n - m$, and the order of the fundamental group of $G.x_0$ equals k.*

Proof. Suppose the isotropy group G_{x_0} has dimension m and k connected components. Let S_{x_0} be a slice at x_0. Then $G.S_{x_0}$ is open. For any $z \in G.S_{x_0}$ we find $z \in g.S_{x_0} = S_{g.x_0}$ (proposition 9.5.1) for some $g \in G$. So $G_z \subseteq G_{g.x_0} = gG_{x_0}g^{-1}$, by proposition 9.4.3(5), and, by the assumptions on the isotropy group G_{x_0}, it follows $G_z = gG_{x_0}g^{-1}$. Since $z \in G.S_{x_0}$ was arbitrary, this means that $G.x_0$ is a principal orbit.

Conversely, if $G.x_0$ is principal, but there were a point $y \in M$ with strictly smaller isotropy group G_y (in the sense of the corollary), then the orbit $G.y$ would be principal, by the previous paragraph, but principal of a different type than $G.x_0$. According to corollary 9.5.6, this is impossible.

Now assume furthermore that G is connected and simply connected. Let us consider the fibration $G_{x_0} \to G \to G/G_{x_0} \cong G.x_0$ and the following part of its long exact homotopy sequence:

$$0 = \pi_1(G) \to \pi_1(G.x_0) \to \pi_0(G_{x_0}) \to \pi_0(G) = 0.$$

Consequently, we have $|\pi_1(G.x_0)| = k$ if and only if the isotropy group G_{x_0} has k connected components. □

9.6. Stratification of the orbit space

This section is dedicated to the study of the natural stratification of the orbit space given by orbit types.

Again we adopt the setting of section 8.2: Let G be a compact Lie group and let $\rho : G \to \mathrm{O}(V)$ be an orthogonal representation in a real finite dimensional Euclidean vector space V with inner product $\langle . \mid . \rangle$. Let $\sigma_1, \dots, \sigma_n$ be a system of homogeneous generators of $\mathbb{R}[V]^G$ with positive degrees, and consider the orbit map $\sigma = (\sigma_1, \dots, \sigma_n) : V \to \mathbb{R}^n$. The following discussion is still correct if one replaces the vector space V by an arbitrary proper Riemannian G-manifold M.

9.6.1. Orbit type submanifolds of V. Let (H) be one particular orbit type ($H = G_v$ for a $v \in V$), see 9.4.1. The union of orbits of type (H), namely

$$V_{(H)} := \bigcup_{(G_x)=(H)} G.x = \{x \in V : (G_x) = (H)\},$$

is called an *orbit type submanifold of the representation* $\rho : G \to \mathrm{O}(V)$, and $V_{(H)}/G = \sigma(V_{(H)})$ is called an *orbit type manifold in the orbit space* $V/G = \sigma(V)$.

CLAIM (1). *$V_{(H)}$ is a smooth G-invariant submanifold of V.*

PROOF. $V_{(H)}$ is of course G-invariant by definition. We only have to prove that it is a smooth submanifold of V. Take any $v \in V_{(H)}$, then without loss of generality $H = G_v$. Let S_v be a slice at v. Consider the tubular neighborhood $G.S_v \cong G \times_H S_v$, see theorem 9.4.4. Then we assert that the orbits of type (H) in $G.S_v$ are just those orbits that meet S_v in the fixed point set S_v^H of H in S_v. Or, equivalently, $(G \times_H S_v)_{(H)} = G \times_H S_v^H$:

($\subseteq$) $[g, s] \in (G \times_H S_v)_{(H)}$ implies $g.s \in (G.S_v)_{(H)}$, i.e., $(H) = (G_{g.s}) = (G_s)$, and, by proposition 9.4.3(5), we have $G_s \subseteq H$. Hence $G_s = H$, by lemma 9.4.1, which means that $s \in S_v^H$, and so $[g, s] \in G \times_H S_v^H$.

($\supseteq$) $[g, s] \in G \times_H S_v^H$ means that $s \in S_v^H$, and in turn $H \subseteq G_s$. On the other hand $G_s \subseteq H$ by proposition 9.4.3(5), therefore $G_s = H$ and so $[g, s] \in (G \times_H S_v)_{(H)}$.

From now on let S_v be a normal slice at v. Since V is a vector space, S_v is simply an open ball centered at 0 in N_v. Let $H = G_v$ act on N_v via the slice representation, then the fixed point set N_v^H is a linear subspace of N_v. Therefore, S_v^H is a submanifold of S_v. Now consider the diagram

$$\begin{tikzcd} G \times S_v^H \arrow[r, "l"] \arrow[d, "q"'] & V \\ G \times_H S_v^H \arrow[ur, "i"'] & \end{tikzcd}$$

The map i is well defined, injective (see the proof of theorem 9.4.4) and smooth, since q is a submersion and l is smooth. Moreover, q is open, and so is l: Consider any open set of the form $U \times W$ in $G \times S_v^H$. Then, $l(U \times W) = \bigcup_{u \in U} l_u(W)$ is open as well, since each l_u is a diffeomorphism. Consequently, i must be open. So i is an embedding, and $G.S_v^H \cong G \times_H S_v^H$ is an embedded submanifold of V. $\square$

In particular, claim 1 yields that $V_{(H)}$ is a proper Riemannian G-manifold, since G is compact and since the restriction of $\langle . \mid . \rangle$ to $V_{(H)}$ defines a G-invariant Riemannian metric on $V_{(H)}$ (again denoted by $\langle . \mid . \rangle$).

9.6.2. Orbit type manifolds in V/G. Let us study the quotient map $\pi : V_{(H)} \to V_{(H)}/G$ and the orbit space $V_{(H)}/G$.

CLAIM (2). *$V_{(H)}/G$ is a smooth manifold.*

PROOF. Let $x \in V_{(H)}$, and let S_x be a normal slice at x (with respect to the action of G on $V_{(H)}$). By proposition 9.4.3(5), we have $G_y \subseteq G_x$ for all $y \in S_x$. Since there is only one orbit type in $V_{(H)}$, G_y must be conjugate to G_x, and both are compact, hence, by lemma 9.4.1, they must be the same. That implies, by lemma 9.5.4, that G_x acts trivially on S_x. From proposition 9.4.3(8) it follows that $\pi(S_x) = G.S_x/G \cong S_x/G_x = S_x$ is an open neighborhood of $\pi(x)$ in $V_{(H)}/G$, and with theorem 9.4.4 we have that $G.S_x$ is isomorphic to $G \times_{G_x} S_x = G/G_x \times S_x$. Therefore, for any $x \in V_{(H)}$, $(\pi(S_x), \exp_x^{-1}|_{S_x})$ can serve as a chart for $V_{(H)}/G$. Obviously, these charts are compatible, whence they form a smooth atlas. By theorem 9.4.7, $V_{(H)}/G$ is Hausdorff, and consequently it is a smooth manifold. $\square$

Now let us consider the quotient map $\pi : V_{(H)} \to V_{(H)}/G$ more carefully. We have seen in the forgoing proof that, for any $x \in V_{(H)}$, $G.S_x \cong G/G_x \times S_x \cong G.x \times S_x$ is a neighborhood of x in $V_{(H)}$ and $\pi(S_x) \cong S_x$ is a neighborhood of $\pi(x)$ in $V_{(H)}/G$. Hence we can identify $T_x V_{(H)} \cong T_x(G.x) \times N_x$ and $T_{\pi(x)}(V_{(H)}/G) \cong N_x$. One finds that π is a smooth submersion.

CLAIM (3). *There exists a Riemannian metric on $V_{(H)}/G$ making the projection $\pi : V_{(H)} \to V_{(H)}/G$ a* Riemannian submersion, *i.e., $T_x\pi : \mathrm{Hor}(\pi)_x := \ker(T_x\pi)^\perp \to T_{\pi(x)}(V_{(H)}/G)$ is an isometric isomorphism for all $x \in V_{(H)}$.*

PROOF. For $X_x, Y_x \in \mathrm{Hor}(\pi)_x = N_x$ we define

$$\gamma_{\pi(x)}(T_x\pi.X_x, T_x\pi.Y_x) := \langle X_x \mid Y_x \rangle_x.$$

This gives a well defined inner product on $T_{\pi(x)}(V_{(H)}/G) \cong N_x$: Choose $X'_{g.x}, Y'_{g.x} \in \mathrm{Hor}(\pi)_{g.x}$ such that $T_{g.x}\pi.X'_{g.x} = T_x\pi.X_x$ and $T_{g.x}\pi.Y'_{g.x} = T_x\pi.Y_x$ (remember that $T\pi$ is surjective). Then, $T_{g.x}\pi.(X'_{g.x} - T_x l_g.X_x) = 0$, so the difference $X'_{g.x} - T_x l_g.X_x$ is vertical, i.e., $X'_{g.x} - T_x l_g.X_x \in \ker(T_{g.x}\pi)$. On the other hand $X'_{g.x}$ is horizontal, and so is $T_x l_g.X_x$, this is because $T l_g$ maps vertical vectors to vertical vectors, since l_g leaves $G.x$ invariant, and, being an isometry, it maps horizontal vectors to horizontal ones. Therefore, $X'_{g.x} - T_x l_g.X_x$ is horizontal as well as vertical and must be zero, i.e., $X'_{g.x} = T_x l_g.X_x$, and in the same way $Y'_{g.x} = T_x l_g.Y_x$. Now we can conclude that

$$\langle X'_{g.x} \mid Y'_{g.x} \rangle_{g.x} = \langle T_x l_g.X_x \mid T_x l_g.Y_x \rangle_{g.x} = \langle X_x \mid Y_x \rangle_x.$$

The Riemannian metric γ on $V_{(H)}/G$ makes π a Riemannian submersion. $\square$

9.6.3. $\pi : V_{(H)} \to V_{(H)}/G$ as associated bundle. Let us finally try to understand in what sense $\pi : V_{(H)} \to V_{(H)}/G$ is an associated bundle. Let $V^H = \{x \in V : H \subseteq G_x\}$ and consider the set

$$V_H := \{x \in V : G_x = H\} = V_{(H)} \cap V^H,$$

where the second equality is a consequence of the definitions and lemma 9.4.1. *We assert that V_H is a geodesically complete submanifold of $V_{(H)}$.* For: Consider first $V^h_{(H)} := \{x \in V_{(H)} : h.x = x\}$ for some $h \in H$. If we choose $X \in T_x V^h_{(H)}$, then $T_x l_h.X = X$ and hence $h.(\exp_x(tX)) = \exp_x(T_x l_h.tX) = \exp_x(tX)$. So the geodesic through x with starting vector X stays in $V^h_{(H)}$. Now $V_H = \bigcap_{h \in H} V^h_{(H)}$, and the assertion follows.

Further, *V_H is $N(H)$-invariant*, where $N(H)$ denotes the normalizer of H in G: $Hn.x = nH.x = n.x$ for $n \in N(H)$ and $x \in V_H$. The restriction $\pi : V_H \to V_{(H)}/G$ is a smooth submersion, since for each $x \in V_H$ the corresponding slice S_x is also contained in V_H, namely, $G_y = G_x = H$ for all $y \in S_x$ as seen before. The fiber of $\pi : V_H \to V_{(H)}/G$ is a free $N(H)/H$-orbit, for if $\pi(x) = \pi(y)$ and $H = G_x = G_y$, then there is a $g \in G$ such that $x = g.y$, whence $gH.y = g.y = x = H.x = Hg.y$ and so $g \in N(H)$. Furthermore, $\pi : V_H \to V_{(H)}/G$ is surjective: Let $\bar{x} \in V_{(H)}/G$ and $x \in V_{(H)}$ such that $\bar{x} = \pi(x)$. There is a $g \in G$ such that $gG_xg^{-1} = H$, and hence $g.x \in V_H$.

So we have proved that $\pi : V_H \to V_{(H)}/G$ is a principal $N(H)/H$-bundle.

CLAIM (4). *$V_{(H)}$ is the associated bundle with fiber G/H.*

PROOF. Consider the following diagram:

$$
\begin{tikzcd}
V_H \times G/H \arrow[d] \arrow[dr, "{(x,[g])\mapsto g.x}"] & \\
V_H \times_{N(H)/H} G/H \arrow[r, "\cong"'] \arrow[d] & V_{(H)} \arrow[d] \\
V_{(H)}/G \arrow[r, equal] & V_{(H)}/G
\end{tikzcd}
$$

which is commutative, since we have

$$
\begin{tikzcd}
S_x \times G/H \arrow[r, "\cong"] \arrow[dr] & \pi^{-1}(S_x) \arrow[d] \\
& S_x
\end{tikzcd}
$$

where S_x is an open neighborhood of x in $V_{(H)}/G$ which lies in V_H. □

In particular, by corollary 9.5.7, the set V_{reg} of regular points in V is exactly the set $V_{(H)}$, where (H) is the minimal orbit type with respect to the ordering defined in section 9.4, namely the principal orbit type. By corollary 9.5.6, the connectedness of V implies that there is precisely one principal orbit type. So, in particular, $\pi : V_{\text{reg}} \to V_{\text{reg}}/G$ is a locally trivial fiber bundle.

9.6.4. Orbit type stratification of V and V/G. *The partition of V in submanifolds $V_{(H)}$ and that of V/G in manifolds $V_{(H)}/G$ is locally finite* which can be seen as follows: We show by induction on the dimension m of V that there occur only finitely many orbit types in V, i.e., the partition of V into $V_{(H)}$ and that of V/G into $V_{(H)}/G$ is even finite. Note that, in general, for a proper G-manifold M and each $x \in M$ there is a G-invariant neighborhood of x in which only finitely many orbit types occur.

For $m = 0$ there is nothing to prove. Suppose the assertion is true for dimensions lower than m. Since the G-action on V is orthogonal, it restricts to an isometric G-action on the unit sphere S^{m-1} in V and makes it a proper Riemannian manifold. Let $x \in S^{m-1}$ and consider a normal slice S_x (with respect to the G-action on S^{m-1}) at x. Then S_x is a Riemannian manifold, and G_x acts isometrically on S_x. Moreover, S_x is equivariantly diffeomorphic to an open ball in $N_x = T_x(G.x)^{\perp}$ (via exp; see definition 9.5.1). So we have $\dim S_x \leq m - 1$, and, thus, by induction hypothesis, S_x has only finitely many G_x-orbit types. By proposition 9.4.3(7), the number of G-orbit types in $G.S_x$ can be no more than the number of G_x-orbit types in S_x. The open sets $G.S_x$ cover S^{m-1} which is compact, hence it has only finitely many orbit types globally. The orbit types are the same on all spheres $r \cdot S^{m-1}$ $(r > 0)$, because $x \mapsto \frac{1}{r}x$ is G-equivariant. Consequently, V has only finitely many orbit types; those of S^{m-1} and the 0-orbit. This completes the proof of the assertion.

For any two closed subgroups $H, K \subseteq G$ with $(H) < (K)$ the intersection $V_{(K)} \cap \overline{V_{(H)}}$ is open and closed in $V_{(K)}$. For: Assume without loss that H is a subgroup of K. By the slice theorem 9.4.4, it suffices to show that $(G \times_K N)_{(K)} = G/K \times N^K \subseteq \overline{(G \times_K N)_{(H)}}$, provided that $(G \times_K N)_{(H)} \neq \emptyset$. We may suppose that the K-module N is endowed with an K-invariant metric, and let W denote the orthogonal space to N^K. Then $(G \times_K N)_{(H)} = G \times_K N_{(H)} = (G \times_K W_{(H)}) \times N^K$. Since $W_{(H)}$ is invariant with respect to multiplication by non-vanishing scalars, the origin of W lies in $\overline{W_{(H)}}$. Hence,

$$(G \times_K N)_{(K)} = G/K \times N^K \subseteq (G \times_K \overline{W_{(H)}}) \times N^K) = \overline{(G \times_K N)_{(H)}},$$

and the claim is proved.

Definition. Let X be a paracompact Hausdorff space with countable topology. Let $\mathcal{Z}$ be a locally finite partition of X into locally closed subspaces $S \subseteq X$. We assume that each $S \in \mathcal{Z}$ is a smooth manifold in the induced topology. The pair $(X, \mathcal{Z})$ is called a *decomposition* of X with *pieces* in $\mathcal{Z}$ when the following condition (*frontier condition*) is satisfied: If $R, S \in \mathcal{Z}$ are such that $R \cap \overline{S} \neq \emptyset$, then $R \subseteq \overline{S}$. In this case we write $R \leq S$. If, in addition, $R \neq S$, we say that R is a boundary piece of S and write $R < S$. The boundary of a locally closed topological subspace A of X is defined as $\partial A := \overline{A} \backslash A$. The boundary ∂S of a piece $S \in \mathcal{Z}$ consists of all pieces $R < S$.

Let $(X, \mathcal{Y})$ and $(X, \mathcal{Z})$ be two decompositions of the same topological space X. We shall say that $\mathcal{Z}$ is *coarser* than $\mathcal{Y}$ or that $\mathcal{Y}$ is *finer* than $\mathcal{Z}$ if for each piece $S \in \mathcal{Y}$ there exists a piece $R_S \in \mathcal{Z}$ such that $S \subseteq R_S$, $\mathrm{id}|_S : S \to R_S$ is smooth, and for all $S \leq S'$ we have $R_S \leq R_{S'}$.

Let X be a topological space and $x \in X$. Two subsets A and B of X are said to be equivalent at x if there exists a neighborhood U of x in X such that $A \cap U = B \cap U$. This defines a equivalence relation on the power set of X. The equivalence class of all sets equivalent at x is called the *set germ* at x.

A *stratification* of the topological space X is a map $\mathcal{S}$ that associates to any $x \in X$ the set germ $\mathcal{S}(x)$ of a closed subset of X such that: For every $x \in X$ there is a neighborhood U of x and a decomposition $\mathcal{Z}$ of U such that for all $y \in U$ the germ $\mathcal{S}(y)$ coincides with the set germ of the piece of $\mathcal{Z}$ that contains y. The pair $(X, \mathcal{S})$ is called a *stratified space*. Any decomposition of X defines a stratification of X by associating to each of its points the set germ of the piece on which it is sitting.

This notion of a stratification is due to Mather [**Mat73**]. In this sense stratifications generate equivalence classes of decompositions of X. One can show that within every such equivalence class there exists a coarsest decomposition, see [**Pfl01**]; its pieces are the *strata* of X.

According to that definition, we have shown that assigning to each point $x \in V$ the set germ $\mathcal{S}(x)$ of the set $V_{(G_x)}$ provides a stratification of V, the *orbit type stratification* of V. Moreover, assigning to each point $G.x$ in the orbit space V/G the germ of the set $V_{(G_x)}/G \cong V_{G_x}/(N(G_x)/G_x)$ we get a stratification of V/G, the *orbit type stratification* of V/G.

Note that this discussion is true in general for proper G-manifolds M instead of the G-module V.

9.6.5. Orbit type and primary stratification of V and V/G coincide. We shall show now that the orbit type stratification of the orbit space $V/G = \sigma(V)$ presented above coincides with its natural stratification as a *semianalyic* subset of $\mathbb{R}^n$; semianalytic means given locally by finitely many analytic equations and inequalities.

A *(primary) stratification* of a semianalytic subset E of $\mathbb{R}^n$ is a locally finite partition of E into connected analytic manifolds, called the *strata*, such that the boundary of each stratum is the union of a set of lower dimensional strata. The natural stratification of a semianalytic subset E of $\mathbb{R}^n$ of dimension p may be constructed in the following manner. Let U^p be the analytic submanifold of those points in E which have a p-dimensional analytic submanifold of $\mathbb{R}^n$ as neighborhood in E; these points are called *regular of dimension* p. We define $U^{p-1}, \dots, U^0$ by decreasing induction as follows: Let $p > q \geq 0$. Put $Z^q := E\backslash(U^p \cup \ldots \cup U^{q+1})$ and denote by W^q the set of regular points of Z^q of dimension q. We define

$$U^q := W^q \cap \bigcap_{j>q} (\operatorname{int}_q(W^q \cap \overline{\Gamma}^j_\nu) \cup \operatorname{int}_q(W^q \backslash \overline{\Gamma}^j_\nu)),$$

where Γ^j_ν are the connected components of U^j, and 'int_q' denotes the interior in W^q. Then $\{\Gamma^j_\nu\}$ is the desired stratification; see [**Bie75**], [**Loj65**]. The following theorem is due to Bierstone [**Bie75**].

THEOREM. *Let G be a compact Lie group and let $\rho : G \to \mathrm{O}(V)$ be an orthogonal representation in a real finite dimensional Euclidean vector space V with inner product $\langle . \mid . \rangle$. Let $\sigma_1, \ldots, \sigma_n$ be a system of homogeneous generators of $\mathbb{R}[V]^G$ with positive degrees, and consider the orbit map $\sigma = (\sigma_1, \ldots, \sigma_n) : V \to \mathbb{R}^n$. Then: The semianalytic (primary) stratification of the orbit space $V/G = \sigma(V)$ coincides with its stratification by components of manifolds of given orbit type.*

PROOF. Let $p = \dim \sigma(V)$, and let X^q be the union of components of dimension q of submanifolds of the orbit space comprising orbits of a given type, i.e., $X^q = \sigma(V_{(H)})$ for a certain type (H). With notation as above we show $X^q = U^q$ for $q = p, p-1, \ldots, 0$, by decreasing induction on q.

Obviously, we have $X^p = \sigma(V_{\mathrm{reg}}) = U^p$. So assume $X^j = U^j$ for $j > q$. It is clear then that $X^q \subseteq W^q$. Consider a point $v \in V$ such that $\sigma(v) \in \sigma(V)\backslash(U^p \cup \ldots \cup U^{q+1} \cup X^q)$. We shall show that then $\sigma(v) \notin W^q$, and hence $W^q \subseteq X^q$. Let S_v be a normal slice at v. The isotropy group G_v acts orthogonally on S_v via the slice representation, and therefore it acts orthogonally on the orthogonal complement T_v in N_v of the fixed point subspace $S_v^{G_v}$. Denoting by $(y_1, \ldots, y_m)$ coordinates in T_v about v, we may choose a set of generators $\psi_1, \ldots, \psi_l$ of the algebra $\mathbb{R}[T_v]^{G_v}$ such that $\psi_1(y) = y_1^2 + \cdots + y_m^2$ and each ψ_i is homogeneous of degree at least 2, since there do not exist non-trivial G_v-invariant linear forms on T_v. Let $\psi = (\psi_1, \ldots, \psi_l)$ be the corresponding orbit map.

By lemma 9.6.6, $\sigma(V)$ is analytically isomorphic near $\sigma(v)$ to a neighborhood of the origin in $S_v^{G_v} \times \psi(T_v) \subseteq S_v^{G_v} \times \mathbb{R}^l$. Now, since $T_v^{G_v} = \{0\}$, the set $\psi(T_v) \subseteq \mathbb{R}^l$ contains no non-singular analytic curves through the origin: Let $c = (c_1, \ldots, c_l)$ be an analytic curve in $\psi(T_v)$ defined near 0 in $\mathbb{R}$ with $c(0) = 0$; then $c_1'(0) = 0$, by the shape of ψ_1, and, hence, $c_i'(0) = 0$ for all i, by the multiplicity lemma 10.1.3. A direct way to see this, is the observation that the image $\psi(T_v)$ lies in $\{(z_1, \ldots, z_l) : z_1 \geq 0, |z_i| \leq D z_1^{\frac{1}{2}\deg \psi_i}, 2 \leq i \leq l\}$, where D is a constant dominating $|\psi_i(y)|$, $1 \leq i \leq l$, for y in the unit sphere of T_v. Hence $\sigma(v) \notin W^q$. So we have proved that $X^q = W^q$.

Using the induction hypothesis, one finds that for each component Γ^j_ν of U^j, $j > q$, we have $\operatorname{int}_q(W^q \backslash \overline{\Gamma}^j_\nu) = W^q \backslash \operatorname{int}_q(W^q \cap \overline{\Gamma}^j_\nu)$:

($\subseteq$) Let $z \in \operatorname{int}_q(W^q \backslash \overline{\Gamma}^j_\nu)$. In particular, $z \notin \overline{\Gamma}^j_\nu$ and so $z \notin \operatorname{int}_q(W^q \cap \overline{\Gamma}^j_\nu)$.

($\supseteq$) Suppose $z \notin \operatorname{int}_q(W^q \backslash \overline{\Gamma}^j_\nu)$, i.e., each open neighborhood of z in W^q contains accumulation points of $\overline{\Gamma}^j_\nu$. We already know that $X^j = U^j$ for $j > q$, and $X^q = W^q$. Therefore, the piece of boundary of Γ^j_ν lying in W^q, namely $W^q \cap \overline{\Gamma}^j_\nu$, must have dimension q. Consequently, there exists a neighborhood of z in W^q consisting entirely of accumulation points of Γ^j_ν; in other words, $z \in \operatorname{int}_q(W^q \cap \overline{\Gamma}^j_\nu)$.

With this identity we see that $U^q = W^q = X^q$, and the theorem is proved. ∎

REMARK. It follows that the orbit type stratification of V/G is a Whitney stratification, i.e., Whitney's condition (b) is satisfied for any pair of strata.

Suppose $(X, \mathcal{S})$ is a stratified space sitting in some $\mathbb{R}^n$. Let $R, S \in \mathcal{S}$. The *Whitney condition* (b) at the point $z \in R$ is given as follows: Let $(x_k)_k \subseteq R$ and $(y_k)_k \subseteq S$ be two sequences with $z = \lim x_k = \lim y_k$ and such that $x_k \neq y_k$, for all k. Suppose that the set of connecting lines $\overline{x_k y_k} \subseteq \mathbb{R}^n$ converges in projective space to a line L and that the sequence of tangent spaces $(T_{y_k} S)_k$ converges in the Grassmann bundle to $\tau \subseteq \mathbb{R}^n$. Then $L \subseteq \tau$. If this condition is verified for every point $z \in R$, the pair (R, S) is said to satisfy the Whitney condition (b). The stratified space $(X, \mathcal{S})$ is called *Whitney stratified* if for every pair of strata Whitney's condition (b) is fulfilled.

In general one can show the following, see [**Pfl01**]: Let M/G be the orbit space of a proper G-manifold M. Then M/G carries a canonical smooth structure the smooth functions of which are given by $C^\infty(\pi^{-1}(U))^G$ for $U \subseteq M/G$ open. By means of this the orbit type stratification of M/G is Whitney and minimal among all Whitney stratifications of M/G.

For detailed information on smooth structures on stratified spaces, Whitney's conditions, etc., we refer to [**Pfl01**].

Remark. If the real vector space V is replaced by a complex vector space, then the semianalytic (primary) stratification of the orbit space is, in general, coarser than the stratification by orbit type.

9.6.6. Lemma. *Consider an orthogonal finite dimensional representation $G \to \mathrm{O}(V)$ of a compact Lie group G, and let $\sigma_1, \ldots, \sigma_n$ be homogeneous generators of positive degree of $\mathbb{R}[V]^G$. Let $v \in V$, $N_v = T_v(G.v)^\perp$, and let S_v be a normal slice at v. Let $\tau_1, \ldots, \tau_m$ be homogeneous generators of positive degree of $\mathbb{R}[N_v]^{G_v}$. Then, the functions $\tau_j|_{S_v}$ are real analytic functions of the $\sigma_i|_{S_v} - \sigma_i|_{S_v}(v)$ in a neighborhood of v. In particular, there is a real analytic isomorphism of a neighborhood of the origin in the orbit space $S_v/G_v = \tau|_{S_v}(S_v) \subseteq \mathbb{R}^m$ with a neighborhood of $\sigma(v)$ in $V/G = \sigma(V)$, where $\sigma = (\sigma_1, \ldots, \sigma_n)$ and $\tau = (\tau_1, \ldots, \tau_m)$.*

Proof. We may assume that $\tau_1, \ldots, \tau_m$ is a *minimal* system of homogeneous generators of positive degree of $\mathbb{R}[N_v]^{G_v}$, i.e., there is no nontrivial polynomial relation of the type

$$\tau_i = P(\tau_1, \ldots, \tau_{i-1}, \tau_{i+1}, \ldots, \tau_m).$$

Let us observe first that the minimality of $\tau_1, \ldots, \tau_m$ is equivalent to the condition that $\tau_1, \ldots, \tau_m$ project to a real vector space basis of I/I^2, where I denotes the ideal in $\mathbb{R}[N_v]^{G_v}$ of polynomials vanishing at the origin. For: To see $(\Leftarrow)$, suppose for contradiction that there is a nontrivial polynomial relation between $\tau_1, \ldots, \tau_m$. It can be written as

$$\tau_i = \sum_{j \neq i} \lambda_j \tau_j + \sum \mu_\alpha \tau^\alpha,$$

where $\lambda_j, \mu_\alpha \in \mathbb{R}$ and the second summation is taken over all multi-indices $\alpha \in \mathbb{N}^m$ with $|\alpha| \geq 2$ and $\alpha_i = 0$. This immediately implies

$$\tau_i \equiv \sum_{j \neq i} \lambda_j \tau_j \mod I^2.$$

So the τ_i are linearly dependent modulo I^2, a contradiction. Conversely, assume $\tau_1, \ldots, \tau_m$ is minimal. Since the τ_i generate $\mathbb{R}[N_v]^{G_v}$, their projections automatically generate I/I^2 as a vector space. Suppose for contradiction that there is a nontrivial relation

$$\sum \lambda_i \tau_i \equiv 0 \mod I^2.$$

Let us order the polynomials τ_i by their degree such that $i < j$ implies $\deg(\tau_i) \leq \deg(\tau_j)$. Let i_0 be the smallest index i for which $\lambda_i \neq 0$. Then, we may express τ_{i_0}

as follows

$$\tau_{i_0} = \sum_{i_0<j} \mu_j \tau_j + \sum_{|\alpha|\geq 2} \nu_\alpha \tau^\alpha.$$

This equality still holds, if we drop all terms of degree $\neq \deg(\tau_{i_0})$. After doing so, we see that τ_{i_0} does not appear on the right-hand side of the equation; if it did, then it would be in a term $\nu_\alpha \tau^\alpha$ with $\alpha_{i_0} \neq 0$ in the sum on the far right, and this term would have degree $> \deg(\tau_{i_0})$. Hence, we have found a nontrivial polynomial relation between the τ_i, in contradiction to their minimality.

Now let us come to the actual proof of the lemma. Let us introduce the notation $\bar{\sigma}_i := \sigma_i|_{S_v}$ and $\bar{\tau}_j := \tau_j|_{S_v}$ for $1 \leq i \leq n$ and $1 \leq j \leq m$. The σ_i are G-invariant and in particular G_v-invariant polynomials, thus there are polynomials $H_1, \dots, H_n$ defined on $\mathbb{R}^m$ such that

$$\bar{\sigma}_i - \bar{\sigma}_i(v) = H_i(\bar{\tau}_1, \dots, \bar{\tau}_m) \quad \text{for } 1 \leq i \leq n. \tag{9.17}$$

In the following, if X is a C^∞ manifold, then $C^\infty(X)$ is meant to be the space of real valued C^∞ functions on X with the topology of uniform convergence of functions and all their derivatives on compact subsets of X, i.e., with the *compact C^∞ topology*. Without loss of generality we may assume that the normal slice S_v was chosen compact. Then, $G.S_v$ is compact as well, and we can apply the Weierstrass approximation theorem to see that polynomials are dense in $C^\infty(G.S_v)$. Averaging over G shows that $\mathbb{R}[V]^G$ is dense in $C^\infty(G.S_v)^G$: Let $f \in C^\infty(G.S_v)^G$ and $p \in \mathbb{R}[V]$ with $\|f - p\|_{G.S_v} \leq \epsilon$, then $q(x) := \int_G p(g.x)dg$ (with dg the Haar measure) is G-invariant and again a polynomial, since the G-action is linear, and we have

$$|f(x) - q(x)| = \left| \int_G f(g.x)dg - \int_G p(g.x)dg \right| \leq \int_G |f(g.x) - p(g.x)| dg \leq \epsilon.$$

We have $G.S_v \cong G \times_{G_v} S_v$ (see theorem 9.4.4), and, hence, if $i : S_v \to G.S_v$ denotes the injection of S_v into $G.S_v$, then

$$i^* : C^\infty(G.S_v)^G \to C^\infty(S_v)^{G_v}$$

is a homeomorphism. Summarizing we see that $(\sigma \circ i)^*\mathbb{R}[\mathbb{R}^n]$ and $(\sigma \circ i)^* C^\infty(\mathbb{R}^n)$ are dense in $C^\infty(S_v)^{G_v}$, where $\sigma = (\sigma_1, \dots, \sigma_n) : V \to \mathbb{R}^n$ (recall here that by the classical theorem of Hilbert and Nagata $\sigma^*\mathbb{R}[\mathbb{R}^n] = \mathbb{R}[V]^G$, e.g. lemma 9.3.3(1)):

$$\begin{aligned} C^\infty(S_v)^{G_v} &= i^* C^\infty(G.S_v)^G = i^* \overline{\mathbb{R}[V]^G} = \overline{i^*\mathbb{R}[V]^G} \\ &= \overline{i^*\sigma^*\mathbb{R}[\mathbb{R}^n]} = \overline{(\sigma \circ i)^*\mathbb{R}[\mathbb{R}^n]} = \overline{(\sigma \circ i)^* C^\infty(\mathbb{R}^n)}. \end{aligned}$$

Now note that $\sigma \circ i = \sigma|_{S_v} = \bar{\sigma}$, hence Glaeser's lemma 9.6.7 guarantees that for $\bar{\sigma} : S_v \to \mathbb{R}^n$, $\bar{\tau}_i \in C^\infty(S_v)^{G_v} = \overline{\bar{\sigma}^* C^\infty(\mathbb{R}^n)}$, and $v \in S_v$ we can find a smooth function $\psi_i \in C^\infty(\mathbb{R}^n)$ such that

$$\mathrm{Tay}_v(\bar{\tau}_i) = \mathrm{Tay}_{\bar{\sigma}(v)}(\psi_i) \circ \mathrm{Tay}_v(\bar{\sigma}),$$

where $\mathrm{Tay}_v(\bar{\tau}_i) \in \mathbb{R}[[x - v]]$ denotes the formal Taylor series of $\bar{\tau}_i$ in v and by the composition on the right-hand side we mean the insertion of $\mathrm{Tay}_v(\bar{\sigma}) \in \mathbb{R}[[x-v]]$ for y in $\mathrm{Tay}_{\bar{\sigma}(v)}(\psi_i) \in \mathbb{R}[[y - \bar{\sigma}(v)]]$. Since $\bar{\tau}_i$ and $\bar{\sigma}$ are polynomials, we can reformulate the above equation to

$$\mathrm{Tay}_v(\bar{\tau}_j) = K_j(\bar{\sigma}_1 - \bar{\sigma}_1(v), \dots, \bar{\sigma}_n - \bar{\sigma}_n(v)) \quad \text{for } 1 \leq j \leq m, \tag{9.18}$$

where the K_j are formal power series with the same coefficients as $\mathrm{Tay}_{\bar{\sigma}(v)}(\psi_i)$. We put $H := (H_1, \dots, H_n) : \mathbb{R}^m \to \mathbb{R}^n$, and let L_j denote the linear part of K_j for $1 \leq j \leq m$. Note that τ_i has no constant term, since it is homogeneous of positive degree. Therefore, (9.17) and (9.18) imply that

$$(L_j \circ H)(\tau_1, \dots, \tau_m) \in \tau_j + I^2 \quad \text{for } 1 \leq j \leq m. \tag{9.19}$$

The minimality of the system $\tau_1, \ldots, \tau_m$ is equivalent to the condition that the $\tau_i + I^2$ form a real vector space basis of I/I^2, as we have pointed out at the beginning of the proof. It follows that we have a well defined algebra isomorphism:

$$I/I^2 \to \mathbb{R}[y_1, \ldots, y_m]_+/\langle y^2 \rangle : \quad \tau_i + I^2 \mapsto [y_i],$$

where $\mathbb{R}[y_1, \ldots, y_m]_+$ denotes the polynomials in $y = (y_1, \ldots, y_m)$ without constant term and $[y_i]$ stands for the class of y_i in $\mathbb{R}[y_1, \ldots, y_m]_+/\langle y^2 \rangle$. Now, if we translate (9.19) to the algebra $\mathbb{R}[y_1, \ldots, y_m]_+/\langle y^2 \rangle$, we obtain

$$(L_j \circ H)(y_1, \ldots, y_m) = y_j + \text{higher order terms} \quad \text{for } 1 \leq j \leq m,$$

for any indeterminates $y_1, \ldots, y_m$. Therefore, the map $H : \mathbb{R}^m \to \mathbb{R}^n$ has an injective differential at the origin. The inverse function theorem and equation (9.17) then show that the $\bar{\tau}_j$ are real analytic functions of the $\bar{\sigma}_i - \bar{\sigma}_i(v)$ in a neighborhood of v.

It is immediate that the map $H : \mathbb{R}^m \to \mathbb{R}^n$ provides a real analytic isomorphism of $\bar{\tau}(S_v)$ and $\sigma(V)$ near the orbit of v. Hence, also the additional statement is proved. □

REMARK. If an algebra is finitely generated, then it automatically possesses a minimal system of generators. We only have to take an arbitrary finite set of generators and recursively drop any elements which can be expressed as polynomials in the others.

9.6.7. LEMMA (Glaeser's lemma [**Gla63a**]). *Consider open subsets $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$, a smooth map $f : U \to V$, and $f^* : C^\infty(V) \to C^\infty(U)$ with the compact C^∞ topology on both spaces. Then, for each $\phi \in \overline{f^* C^\infty(V)}$ and for all $a \in U$, there is a $\psi \in C^\infty(V)$ such that*

$$\operatorname{Tay}_a(\phi) = \operatorname{Tay}_{f(a)}(\psi) \circ \operatorname{Tay}_a(f),$$

where $\operatorname{Tay}_a(\phi) \in \mathbb{R}[[x-a]]$ denotes the formal Taylor series of ϕ in a and by the composition on the right-hand side we mean the insertion of $\operatorname{Tay}_a(f) \in \mathbb{R}[[x-a]]$ for y in $\operatorname{Tay}_{f(a)}(\psi) \in \mathbb{R}[[y - f(a)]]$.

PROOF. The assertion of the lemma is equivalent to the statement

$$\operatorname{Tay}_a(f^* C^\infty(V)) = \operatorname{Tay}_a(\overline{f^* C^\infty(V)}),$$

since $\operatorname{Tay}_a(f^* C^\infty(V))$ is simply the set of all jets which can be written as a composition as in the lemma. Due to the fact that Tay_a is continuous, we have the following inclusions:

$$\operatorname{Tay}_a(f^* C^\infty(V)) \subseteq \operatorname{Tay}_a(\overline{f^* C^\infty(V)}) \subseteq \overline{\operatorname{Tay}_a(f^* C^\infty(V))}.$$

Therefore, it is sufficient to show that $\operatorname{Tay}_a \circ f^*$ has a closed image. Since $C^\infty(V)$ is a reflexive Fréchet space, we can show instead that the dual map $(\operatorname{Tay}_a \circ f^*)'$ has a closed image. Now,

$$(\operatorname{Tay}_a)' : \mathbb{R}[[x-a]]' \to C^\infty(U)',$$

where $\mathbb{R}[[x-a]]'$ is the space of all distributions with support $\{a\}$. Let $\sum_\beta \lambda_\beta \delta_a^{(\beta)}$ be such a distribution, and take any $\alpha \in C^\infty(V)$. Then, we have

$$\begin{aligned}\langle \alpha, (\operatorname{Tay}_a \circ f^*)' \sum_\beta \lambda_\beta \delta_a^{(\beta)} \rangle &= \langle (\operatorname{Tay}_a \circ f^*)(\alpha), \sum_\beta \lambda_\beta \delta_a^{(\beta)} \rangle \\ &= \sum_\beta \lambda_\beta (\alpha \circ f)^{(\beta)}(a) = \sum_\gamma \mu_\gamma \partial^{(\gamma)} \alpha(f(a)) = \langle \alpha, \sum_\gamma \mu_\gamma \delta_{f(a)}^{(\gamma)} \rangle.\end{aligned}$$

It follows that the image of $\mathbb{R}[[x-a]]'$ under $(\operatorname{Tay}_a \circ f^*)'$ is contained in the space of all distributions concentrated at $f(a)$ which is isomorphic to a countable sum of $\mathbb{R}$

with the finest locally convex topology. But in this topology every linear subspace is closed (since every quotient mapping is continuous), hence $(\mathrm{Tay}_a \circ f^*)'(\mathbb{R}[[x-a]]')$ is closed. The proof is complete. $\square$

9.6.8. As a consequence of theorem 9.5.5, and, since $T_{\pi(v)}(V_{(H)}/G) \cong N_v^{G_v}$, we can compute the dimension of the stratum $V_{(H)}/G$ of the orbit space of type $(H) = (G_v)$ as follows:

$$\dim V_{(H)}/G = \dim N_v^{G_v} = \operatorname{rank} d\sigma(v) = \operatorname{rank} B(v) = \operatorname{rank} \tilde{B}(\sigma(v)),$$

where B, respectively $\tilde{B}$, denotes the generalized Bezoutiant defined in section 9.2.

Finally, note that, as seen in the proof of theorem 9.6.5, the stratification of $\sigma(V) = V/G$ in a neighborhood of each $\sigma(v)$ is naturally isomorphic to the stratification of N_v/G_v in a neighborhood of 0.

9.6.9. Orbit type stratification of $\mathbb{R}_n/\mathrm{S}_n$. To conclude this section let us investigate the stratification of the orbit space in the case when the symmetric group S_n acts on $\mathbb{R}^n$.

The orbit type stratification of $\mathbb{R}^n$ as S_n-space is determined completely as for every finite reflection group by its reflections and thus by its roots. More precisely: A partition $(\lambda_1, \dots, \lambda_t)$ of the number n is an unordered tuple of positive integers such that $n = \lambda_1 + \dots + \lambda_t$. For $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ consider its multiplicity vector given by the unordered tuple of the multiplicities of coordinates appearing in y. The set of all points y whose multiplicity vector agrees with some fixed partition $(\lambda_1, \dots, \lambda_t)$ of n constitutes a t-dimensional orbit type stratum of $\mathbb{R}^n$. This defines a stratification of $K = \{y \in \mathbb{R}^n : y_1 \leq \dots \leq y_n\}$, the closure of a fundamental domain, too. Then K is homeomorphic to the orbit space $\mathbb{R}^n/S_n$, and, the stratification of K coincides with the orbit type stratification of $\mathbb{R}^n/S_n$, via this identification.

CHAPTER 10

Lifting curves over invariants real analytically and smoothly

This chapter presents many results concerning our lifting problem (see section 8.1 and section 8.2) for real analytic and smooth curves in the orbit space. It is based on [**AKLM00**].

10.1. Local lifting

Let G be a compact Lie group and let $\rho : G \to \mathrm{O}(V)$ be an orthogonal representation in a real finite dimensional Euclidean vector space V with inner product $\langle . \mid . \rangle$. Let $\sigma_1, \dots, \sigma_n$ be a system of homogeneous generators of $\mathbb{R}[V]^G$ with positive degrees $d_1, \dots, d_n$, and consider the orbit map $\sigma = (\sigma_1, \dots, \sigma_n) : V \to \mathbb{R}^n$.

Similarly as in the case of choosing roots of polynomials smoothly, see section 4.2, we shall construct an algorithm which solves the lifting problem locally.

10.1.1. Lifting at regular orbits. We investigate at first the lifting at regular orbits. This corresponds to lemma 4.2.2. By an *orthogonal lift* we mean a lift $\bar{c} : \mathbb{R} \to V$ (at least differentiable) meeting orbits orthogonally, i.e., $\langle \bar{c}'(t) \mid T_{\bar{c}(t)}(G.\bar{c}(t)) \rangle = 0$ for all $t \in \mathbb{R}$.

Lemma. *A smooth (real analytic) curve $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ admits a smooth (real analytic) orthogonal lift $\bar{c}$ in a neighborhood of a regular point $c(t_0) \in V_{\mathrm{reg}}/G$. It is unique up to a transformation from G.*

Proof. The orthogonal distribution $V_{\mathrm{reg}} \ni v \mapsto N_v$ of the locally trivial fiber bundle $\pi : V_{\mathrm{reg}} \to V_{\mathrm{reg}}/G$ defines a real analytic Ehresmann connection in π. A local orthogonal lift of the curve c is the same as a horizontal lift with respect to this connection, near t_0. It is given uniquely by its initial value. See [**KMS93**, Section 9]. □

10.1.2. Fixed points. Consider the subspace V^G of G-invariant vectors in V, and let V' be its orthogonal complement in V. Then $V = V^G \oplus V'$, $V/G = V^G \times V'/G$, and the canonical bilinear map $\mathbb{R}[V^G] \times \mathbb{R}[V']^G \to \mathbb{R}[V]^G$ induces an isomorphism $\mathbb{R}[V]^G \cong \mathbb{R}[V^G] \otimes \mathbb{R}[V']^G$. In this situation the following lemma is obvious:

Lemma. *Any lift $\bar{c}$ of a curve $c = (c_0, c_1)$ of class C^k $(k = 0, 1, \dots, \infty, \omega)$ in $V^G \times V'/G$ has the form $\bar{c} = (c_0, \bar{c}_1)$, where $\bar{c}_1$ is a lift of c_1 to V' of class C^k. The lift $\bar{c}$ is orthogonal if and only if the lift $\bar{c}_1$ is orthogonal.* □

10.1.3. Multiplicity. Remind of the definition of the multiplicity or order of flatness of a continuous function f defined near 0 in $\mathbb{R}$, given in definition 4.2.4:

$$m(f) := \sup\{p \in \mathbb{Z} : f(t) = t^p g(t) \text{ near } 0 \text{ for continuous } g\}.$$

Let $c = (c_1, \dots, c_n)$ be a smooth curve in $\sigma(V) \subseteq \mathbb{R}^n$ with $c(0) = 0$. By possibly increasing the number of generators $\sigma_1, \dots, \sigma_n$ of $\mathbb{R}[V]^G$, we may assume from now on without loss that $\sigma_1 : v \mapsto \langle v \mid v \rangle$ is the Euclidean metric. Then, we have $c_1(t) \geq 0$ for all $t \in \mathbb{R}$, and consequently, $m(c_1) = 2r > 0$, where $r \in \mathbb{N}$ or $r = \infty$.

LEMMA (Multiplicity lemma). *In this situation we have $m(c_i) \geq d_i r$, for all $1 \leq i \leq n$. Remember $d_1, \dots, d_n$ are the degrees of homogeneity of the generators $\sigma_1, \dots, \sigma_n$.*

PROOF. For contradiction suppose that for some $k \geq 2$ we have $m(c_k) < d_k r$. Then $m := \min\{m(c_1)/d_1, \cdots, m(c_n)/d_n\} < r$. We consider the following continuous curve in $\mathbb{R}^n$ for $t \geq 0$:

$$c_{(m)}(t) := (t^{-2m} c_1(t), t^{-d_2 m} c_2(t), \dots, t^{-d_n m} c_n(t)).$$

By the choice of the generators $\sigma_1, \dots, \sigma_n$, we find that $c_{(m)}(t) \in \sigma(V)$ for $t > 0$, and since $\sigma(V)$ is closed in $\mathbb{R}^n$, by its explicit description in theorem 9.2.2, also $c_{(m)}(0) \in \sigma(V)$. Since $m < r$, the first coordinate of $c_{(m)}(t)$ vanishes at $t = 0$. Then $\sigma^{-1}(c_{(m)}(0)) = \{0\}$ and therefore $c_{(m)}(0) = 0$, again since σ_1 is the squared norm on V. In particular, for those j with $m(c_j) = d_j m$ we get a contradiction. □

10.1.4. If $r < \infty$, we shall consider the following smooth curve in $\sigma(V)$:

$$c_{(r)}(t) := (t^{-2r} c_1(t), t^{-d_2 r} c_2(t), \dots, t^{-d_n r} c_n(t)). \tag{10.1}$$

This curve will be useful to reduce the lifting problem in the following sense: We have $c_{(r)}(0) \neq 0$, since $m(c_1) = 2r$. If $c_{(r)}$ is liftable at 0 and $\bar{c}_{(r)}$ is its smooth (real analytic) lift, then $\bar{c}(t) := t^r \cdot \bar{c}_{(r)}(t)$ is a smooth (real analytic) lift of c near 0. If $\bar{c}_{(r)}$ is an orthogonal lift, then also $\bar{c}$, and conversely, since the action of G commutes with homotheties of V. Moreover, the orthogonal lift of c is uniquely determined up to the action of a constant element in G if and only if the orthogonal lift of $c_{(r)}$ has this property.

10.1.5. Local real analytic lifting. After this preliminary work we can attack the local lifting problem for real analytic curves.

THEOREM. *Let $c = (c_1, \dots, c_n) : \mathbb{R} \to \sigma(V) \subseteq \mathbb{R}^n$ be a real analytic curve. Then there exists a real analytic lift $\bar{c}$ in V of c, locally near each $t_0 \in \mathbb{R}$.*

PROOF. Without loss of generality we may assume that $t_0 = 0$. We shall show that there exist local real analytic lifts of c through any $v \in \sigma^{-1}(c(0))$. The proof is carried out by the following algorithm in four steps which generalizes the algorithm 4.2.9.

(1) If $c(0) \neq 0$ corresponds to a regular orbit, unique local orthogonal real analytic lifts exist through all $v \in \sigma^{-1}(c(0))$, by lemma 10.1.1.
(2) If $V^G \neq \{0\}$, then we remove fixed points, by lemma 10.1.2. That lowers the dimension of the vector space under observation.
(3) If $V^G = \{0\}$ and $c(0) \neq 0$ corresponds to a singular orbit, then to each $v \in \sigma^{-1}(c(0))$ we consider the respective slice representation $G_v \to \mathrm{O}(N_v)$. By theorem 9.5.3, the lifting problem reduces to the same problem in N_v/G_v, where the curve is now passing through 0. Note that G_v is smaller than G, since $v \neq 0$ and $V^G = \{0\}$. If $N_v^{G_v} \neq \{0\}$, we continue in step (2). If $N_v^{G_v} = \{0\}$, then continue in step (4).
(4) If $V^G = \{0\}$ and $c(0) = 0$, then $m(c_1) = 2r$ for some $r \in \mathbb{N}$ or $r = \infty$. In the latter case $c_1 = 0$, since c_1 is real analytic. This implies that $c = 0$ is constant which clearly can be lifted. In the former case, by the multiplicity lemma 10.1.3, we have $m(c_i) \geq d_i r$ for all i, and the lifting problem reduces to the curve $c_{(r)}$ defined in equation (10.1). Then $c_{(r)}(0) \neq 0$, and we may continue in the steps (1), (2), or (3).

This algorithm always stops, since each step either gives a local lift, or reduces the lifting problem to a smaller group or a smaller space (see remark (2) after the proof). This completes the proof. □

REMARKS. (1) Note that the role of the splitting lemma 4.2.3 in part 1 is now played by the transition to the slice representation provided by theorem 9.5.3.

(2) When we speak of smaller spaces here we intend lower dimensional vector spaces, of course. In the case of groups we mean it in the following sense: For compact G' and G we write $G' < G$ and say that G' is smaller than G, if

- $\dim G' < \dim G$ or
- if $\dim G' = \dim G$, then G' has less connected components than G.

(3) Note that the case treated in step (4), when $c(0) = 0$, has to be considered separately, since $G.0 = \{0\}$ whence $G_0 = G$ and $N_0 = V$. That is why at 0 we do not gain anything by passing to the slice representation, and so 0 can be considered the 'most' singular point.

10.1.6. Normal nonflatness. Theorem 10.1.5 solves our problem locally for real analytic curves in the orbit space. Now we try to tackle the problem for smooth curves in $\sigma(V)$. As seen in section 4.2 in the special case of S_n acting on $\mathbb{R}^n$, this will not be possible in full generality. Remember that there we had to impose certain genericity conditions: No two roots should meet of infinite order. Let us try to formulate the appropriate genericity conditions also in the general setting. The point here is that, in the smooth case, the algorithm in the proof of theorem 10.1.5 fails to work in only one particular place: In step (4) we can not follow from $r = \infty$ that c_1 vanishes identically. So, when we formulate the extra conditions for the smooth curve c in the orbit space, we have to take care that this implication remains valid.

DEFINITION. Let $s \in \mathbb{N}_0$. Denote by A_s the union of all strata X of the orbit space V/G with $\dim X \leq s$, and by I_s the ideal of $\mathbb{R}[V/\!\!/G] = \mathbb{R}[V]^G$ consisting of all polynomials vanishing on A_{s-1}.

Let $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a smooth curve, $t_0 \in \mathbb{R}$, and $s = s(c, t_0)$ a minimal integer such that, for a neighborhood J of t_0 in $\mathbb{R}$, we have $c(J) \subseteq A_s$. The curve c is called *normally nonflat at* t_0, if there is a $f \in I_s$ such that $f \circ c$ is nonflat at t_0, i.e., the Taylor series of $f \circ c$ at t_0 is not identically zero. This automatically holds if $c(t_0) \notin A_{s-1}$.

A smooth curve $c : \mathbb{R} \to V/G = \sigma(V) \in \mathbb{R}^n$ is called *generic*, if c is normally nonflat at all $t \in \mathbb{R}$. A real analytic curve is automatically generic.

Now we have to clarify, whether the notion of normally nonflatness is invariant under the reduction process used in the proof of theorem 10.1.5.

PROPOSITION. *If a smooth curve $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ is normally nonflat at $t_0 \in \mathbb{R}$, then curves which we obtain from the above reduction process, i.e., removing of fixed points, passing to slice representations, or replacing c by $c_{(r)}$ (see equation* (10.1)*), are normally nonflat at t_0 as well.*

PROOF. Removing fixed points: Suppose $V^G \neq \{0\}$ and let $\dim V^G = k$. In the notation introduced in 10.1.2, each stratum X of $V/G = V^G \times V'/G$ has the form $V^G \times X_1$, where X_1 is a stratum of V'/G. Let $c = (c_0, c_1)$ be a smooth curve in $V/G = V^G \times V'/G$. Suppose $f \in I_s \subseteq \mathbb{R}[V]^G$ is a function such that $f \circ c$ is nonflat at t_0. Since $\mathbb{R}[V]^G = \mathbb{R}[V^G] \otimes \mathbb{R}[V']^G$, we can write $f = \sum_i \phi_i \otimes f_i$, where $\phi_i \in \mathbb{R}[V^G]$ and $f_i \in I'_{s-k}$, the ideal consisting of all polynomials vanishing on all strata of V'/G of dimension strictly lower than $s-k$. Moreover, we have that $f_i \circ c_1$ is nonflat at t_0 for some i. That is, c_1 is normally nonflat at t_0.

Passing to slice representations: If $V^G = \{0\}$ and $c(t_0) \neq 0$, then the statement of the proposition follows from the observation that the stratification of V/G is locally isomorphic to the stratification of N_v/G_v near 0 (see section 9.6) and from theorem 9.5.5, since the notion of normal nonflatness is local.

Replacing c by $c_{(r)}$: Let $V^G = \{0\}$, $c(t_0) = 0$, $s = s(c, t_0)$ minimal such that $c(J) \subseteq A_s$ for a neighborhood J of t_0, and $f \in I_s$ be such that $f \circ c$ is nonflat at t_0. Without loss we can assume that $t_0 = 0$ and that f is homogeneous. Then the function $f \circ c_{(r)}$ is nonflat at 0. □

The following theorem gives the best practical way to check the normal nonflatness of a curve c, in terms of the minors $\tilde{\Delta}^{j_1,\dots,j_k}_{i_1,\dots,i_k}$ of $\tilde{B}$, see section 9.2.

Theorem. *Let $c : \mathbb{R} \to \sigma(V) \subseteq \mathbb{R}^n$ be a smooth curve. Then, c is normally nonflat at $t_0 \in \mathbb{R}$, if the following two conditions are satisfied for some $1 \leq r \leq n$:*
(1) *The functions $\tilde{\Delta}^{j_1,\dots,j_k}_{i_1,\dots,i_k} \circ c$ vanish in a neighborhood of t_0 whenever $k > r$.*
(2) *There exists a minor $\tilde{\Delta}^{j_1,\dots,j_r}_{i_1,\dots,i_r}$ such that $\tilde{\Delta}^{j_1,\dots,j_r}_{i_1,\dots,i_r} \circ c$ is nonflat at t_0.*

Proof. Let $s = s(c, t_0)$ be minimal such that $c(J) \subseteq A_s$ for a neighborhood J of t_0. Since the dimension of the stratum of type (G_v) equals the rank of $\tilde{B}(\sigma(v))$, see 9.6.8, the conditions of the theorem imply $r = s$. Moreover, $\tilde{\Delta}^{j_1,\dots,j_r}_{i_1,\dots,i_r} \in \mathbb{R}[V/\!\!/G] = \mathbb{R}[V]^G$, and it vanishes on A_{r-1}, by the same argumentation, i.e., $\tilde{\Delta}^{j_1,\dots,j_r}_{i_1,\dots,i_r} \in I_r$. But that just means that c is normally nonflat at t_0. □

10.1.7. Local smooth lifts. With these ingredients we can attack the problem of lifting smooth curves locally:

Theorem. *Let $c = (c_1, \dots, c_n) : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a smooth curve which is normally nonflat at $t_0 \in \mathbb{R}$. Then there exists a smooth lift $\bar{c}$ in V of c, locally near t_0.*

Proof. The proof is the same as the one of theorem 10.1.5, since by proposition 10.1.6 the normal nonflatness remains invariant under the reduction process and it guarantees that in step (4) from $r = \infty$ follows $c_1 = 0$. □

10.1.8. Uniqueness of local smooth lifts.

Lemma. *Let $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a smooth curve which is normally nonflat at $t_0 \in \mathbb{R}$. Suppose that $\bar{c}_1$ and $\bar{c}_2$ are smooth lifts in V of c on an open interval I containing t_0. Then there exists a smooth curve g in G defined near t_0 such that $\bar{c}_1(t) = g(t).\bar{c}_2(t)$ for all t near t_0. The real analytic version of this result is also true.*

Proof. The proof follows the algorithm in the proof of theorem 10.1.5.

Without loss of generality let $t_0 = 0$, and we can assume $\bar{c}_1(0) = \bar{c}_2(0) =: v$, by applying a transformation of G to, say, $\bar{c}_2$ if necessary. For a normal slice S_v at v we know that $p : G.S_v \cong G \times_{G_v} S_v \to G/G_v \cong G.v$ is the projection of a fiber bundle associated to the principal bundle $\pi : G \to G/G_v$. Then $p \circ \bar{c}_1$ and $p \circ \bar{c}_2$ are two smooth curves in G/G_v defined near $t = 0$ which admit smooth lifts g_1 and g_2 into G with $g_1(0) = g_2(0) = e$, via the horizontal lift of a principal connection, say. Consequently, $t \mapsto g_j(t)^{-1}.\bar{c}_j(t)$ $(j = 1, 2)$ are two smooth curves in S_v and lifts of c:

$$p(g_j(t)^{-1}.\bar{c}_j(t)) = g_j(t)^{-1}p(\bar{c}_j(t)) = g_j(t)^{-1}\pi(g_j(t)) = g_j(t)^{-1}g_j(t).v = v.$$

This reduces the problem to the group G_v acting on N_v. If v is a regular point, then this action is trivial, and these lifts are automatically the same, so we are done.

If instead v is a singular point and $N_v^{G_v} \neq \{0\}$, we remove the nontrivial fixed points, by lemma 10.1.2. Thus, we may assume that $c(0) = 0$ and $V^G = \{0\}$. In the case that c vanishes identically, the statement is trivial. So we can suppose that the first component of c has multiplicity $2r < \infty$, since c is normally nonflat at 0 by assumption. Then, $t^{-r}\bar{c}_1(t)$ and $t^{-r}\bar{c}_2(t)$ are smooth lifts of $c_{(r)}$, defined in equation (10.1). If we can find a smooth curve $g(t) \in G$ taking $t^{-r}\bar{c}_2(t)$ to $t^{-r}\bar{c}_1(t)$, then we also have $g(t).\bar{c}_2(t) = \bar{c}_1(t)$. The two lifts $t^{-r}\bar{c}_1(t)$ and $t^{-r}\bar{c}_2(t)$ of $c_{(r)}$ can then be fed again into the algorithm.

In the real analytic situation the proof is the same. □

10.2. Global lifting

Here we shall glue together the local smooth lifts found in the previous section in order to get a global smooth lift.

10.2.1. Theorem. *Let $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a generic smooth curve. Then there exists a global smooth lift $\bar{c} : \mathbb{R} \to V$ of c.*

Proof. By theorem 10.1.7, there exist local smooth lifts of c near any $t \in \mathbb{R}$. It is sufficient to prove that each local smooth lift of c defined on an open interval I can be extended smoothly to a larger interval whenever $I \neq \mathbb{R}$.

Suppose $\bar{c}_1 : I \to V$ is a local smooth lift of c, and suppose the open interval I is bounded from above, say, and t_0 is its upper boundary point. By theorem 10.1.7, there exists a local smooth lift $\bar{c}_2$ of c near t_0, and there is a $t_1 < t_0$ such that both $\bar{c}_1$ and $\bar{c}_2$ are defined near t_1. Then lemma 10.1.8 provides the existence of a smooth curve g in G, locally defined near t_1, such that $\bar{c}_1(t) = g(t).\bar{c}_2(t)$. We consider the right logarithmic derivative $X(t) = T_{g(t)}(\mu^{g(t)^{-1}}).g'(t) = g'(t).g(t)^{-1} \in \mathfrak{g} = \mathrm{Lie}(G)$, where $\mu(h,g) = \mu_h(g) = \mu^g(h) = hg$ denotes the multiplication in G. Choose a smooth function $\chi(t)$ which is 1 for $t \le t_1$ and becomes 0 before g ceases to exist. Consequently, $Y(t) = \chi(t)X(t)$ is a smooth curve in $\mathfrak{g}$ defined near $[t_1, \infty)$. The differential equation $h'(t) = Y(t).h(t)$ with initial condition $h(t_1) = g(t_1)$ then has a solution h in G defined near $[t_1, \infty)$ which coincides with g below t_1. Therefore,

$$\bar{c}_{12}(t) := \begin{cases} \bar{c}_1(t) & \text{for} \quad t \le t_1 \\ h(t).\bar{c}_2(t) & \text{for} \quad t \ge t_1 \end{cases}$$

is a smooth lift of c defined on on a larger interval than $\bar{c}_1$. This completes the proof. □

Note that this proof does not work in the real analytic case, since in generality we will not find a real analytic function χ with the required properties because of the lack of C^ω cutoff functions.

10.3. Polar representations

In this section we show that, if we restrict to a subclass of orthogonal representations of compact Lie groups, then we can achieve global orthogonal real analytic or smooth lifts which are unique up to the action of a constant element in G. Recall that by an orthogonal lift we mean a lift meeting orbits orthogonally.

The mentioned subclass of representations is the class of polar representations:

10.3.1. Polar representations.

Definition. An orthogonal representation $\rho : G \to \mathrm{O}(V)$ of a Lie group G on a finite dimensional Euclidean vector space V is called *polar*, if there exists a linear subspace $\Sigma \subseteq V$, called a *section* or a *Cartan subspace*, which meets each orbit orthogonally. See [**Dad85**], [**DK85**], and [**PT87**].

Examples. (1) Every orthogonal representation of a finite group G on V is polar with section $\Sigma = V$.

(2) The standard action of $\mathrm{O}(n)$ on $\mathbb{R}^n$ is polar. Each 1-dimensional linear subspace is a section.

(3) Let $\mathrm{S}(n)$ be the space of real symmetric $n \times n$ matrices. Then the action $\mathrm{O}(n) \times \mathrm{S}(n) \to \mathrm{S}(n) : (A, B) \mapsto ABA^{-1} = ABA^\top$ is polar, where the space Σ of all real diagonal matrices is a section. In fact, that Σ meets every $\mathrm{O}(n)$-orbit is clear from linear algebra, and Σ intersects each orbit orthogonally in terms of the inner product $\langle A \mid B\rangle = \operatorname{trace}(AB^\top) = \operatorname{trace}(AB)$: Let $A \in \Sigma$. For any $X \in \mathfrak{o}(n)$ let ζ_X denote the corresponding fundamental vector field on $\mathrm{S}(n)$. Then

$$\zeta_X(A) = \frac{d}{dt}\Big|_{t=0} \exp(tX) A \exp(tX^\top) = XA + AX^\top = XA - AX.$$

The inner product with $Y \in T_A\Sigma \cong \Sigma$ computes to

$$\begin{aligned}\langle \zeta_X(A) \mid Y\rangle &= \operatorname{trace}(\zeta_X(A)Y) = \operatorname{trace}((XA - AX)Y) \\ &= \operatorname{trace}(XAY) - \operatorname{trace}(AXY) = \operatorname{trace}(XYA) - \operatorname{trace}(XYA) = 0.\end{aligned}$$

Remark. Note that in the more general situation of a connected complete Riemannian G-manifold M with *effective* G-action (i.e. $g.x = x$ for all $x \in M$ implies $g = e$) a *section* is defined to be a connected closed complete submanifold $\Sigma \subseteq M$ which meets each orbit orthogonally. See [**Mic97**], [**PT88**].

10.3.2. The generalized Weyl group. Suppose we are given a polar representation $\rho : G \to \mathrm{O}(V)$ of a compact Lie group G on a finite dimensional Euclidean vector space V, and let Σ be a section. We consider the largest subgroup of G which induces an action on Σ,

$$N(\Sigma) := \{g \in G : l_g(\Sigma) = \Sigma\},$$

and the subgroup of $N(\Sigma)$ consisting of all $g \in G$ which act trivially on Σ,

$$Z(\Sigma) := \{g \in G : l_g(s) = s \text{ for all } s \in \Sigma\}.$$

Since Σ is closed, so is $N(\Sigma)$, and hence it is a Lie subgroup of G. The subgroup $Z(\Sigma) = \bigcap_{s\in\Sigma} G_s$ is closed as well, and it is normal in $N(\Sigma)$. Therefore, $N(\Sigma)/Z(\Sigma)$ is a Lie group, and it acts on Σ effectively. This group is called the *generalized Weyl group of* Σ and is denoted by

$$W(\Sigma) = N(\Sigma)/Z(\Sigma).$$

The group $W(\Sigma)$ *is a finite:* Take a regular point $v \in \Sigma$ and consider a normal slice S_v at v. Then $S_v \subseteq \Sigma$ is open. Hence, any $g \in N(\Sigma)$ close to the identity element maps v back into S_v. By proposition 9.4.3(2), we have $g \in G_v$. Now $G_v = Z(\Sigma)$, since v is regular and so G_v acts trivially on Σ, whence $G_v \subseteq Z(\Sigma)$; the inverse inclusion is obvious. That means that $Z(\Sigma)$ is an open subset of $N(\Sigma)$, and, consequently, the quotient $W(\Sigma)$ is discrete. Since G is compact, $W(\Sigma)$ has to be finite.

Moreover, the generalized Weyl group $W(\Sigma)$ is a reflection group if G is connected, see [**DK85**].

We shall need the following generalization of Chevalley's restriction theorem, which is due to Dadok and Kac and independently to Terng (with more general assumptions than presented here). The proof is omitted here.

Theorem ([**DK85**], [**Ter85**]). *Let $\rho : G \to \mathrm{O}(V)$ be a polar representation of a compact Lie Group, with section Σ and generalized Weyl group $W(\Sigma)$. Then the algebra $\mathbb{R}[V]^G$ of G-invariant polynomials on V is isomorphic to the algebra $\mathbb{R}[\Sigma]^{W(\Sigma)}$ of $W(\Sigma)$-invariant polynomials on the section Σ, via restriction $f \mapsto f|_\Sigma$.*

As a consequence of this theorem we obtain that the orbit spaces $V/G = \sigma(V)$ and $\Sigma/W(\Sigma) = \sigma|_\Sigma(\Sigma)$ are isomorphic as stratified spaces.

10.3.3. Polar representations admit orthogonal lifts.

THEOREM. *Let $\rho : G \to \mathrm{O}(V)$ be a polar representation of a compact Lie group on a finite dimensional Euclidean vector space with orbit map $\sigma : V \to \mathbb{R}^n$. Let $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a curve in the orbit space which is either real analytic or smooth but generic. Then there exists a global orthogonal real analytic or smooth lift $\bar{c} : \mathbb{R} \to V$ which is unique up to the action of a constant element in G.*

PROOF. Let Σ be a section. By theorem 10.3.2, $\sigma|_\Sigma : \Sigma \to \mathbb{R}^n$ is the orbit map for the representation $W(\Sigma) \to \mathrm{O}(\Sigma)$. If c is a generic smooth curve in $\sigma(V) \cong \sigma|_\Sigma(\Sigma)$, then by theorem 10.2.1 there exists a global smooth lift $\bar{c} : \mathbb{R} \to \Sigma$, which as a curve in V is orthogonal to each G-orbit it meets, by the properties of Σ. Note for further use that $\bar{c}$ is nowhere flat, since otherwise the curve c is not generic at some t, which can easily been seen from theorem 10.1.6.

If c is real analytic, there are local real analytic lifts over $\sigma|_\Sigma$ into Σ by theorem 10.1.5. By lemma 10.1.8, these local lifts are unique up to the action of a constant element in $W(\Sigma)$, since $W(\Sigma)$ is finite. Thus we can glue the local lifts to a global real analytic lift $\bar{c}$ in Σ, which as curve in V is an orthogonal lift.

It remains to show that for two global orthogonal lifts $\bar{c}_1, \bar{c}_2 : \mathbb{R} \to V$ of c, there is a constant element $g \in G$ such that $\bar{c}_1(t) = g.\bar{c}_2(t)$ for all t. We may assume that $\bar{c}_1$ lies in a section Σ, by the considerations at the beginning of the proof.

Since c is generic, $\bar{c}_1$ meets each stratum of V only in isolated points, if it is not entirely contained in this stratum. Let $v \in \Sigma$ be arbitrary, then $\Sigma \subseteq N_v = T_v(G.v)^\perp$, and so for the points x in $\Sigma \cap S_v$, which is a neighborhood of v in Σ, we have $G_x \subseteq G_v$, by proposition 9.4.3(5). From these two observations it follows that for an open dense subset $J \subseteq \mathbb{R}$ the groups $G_{\bar{c}_1(t)}$ all agree for $t \in J$ (by lemma 9.4.1), call them H, and we have $H \subseteq G_{\bar{c}_1(t)}$ for all $t \in \mathbb{R}$.

From lemma 10.1.8 we get that $\bar{c}_1(t) = g(t).\bar{c}_2(t)$ for some smooth curve $g : I \to G$, locally near each t_0. Let us consider the right logarithmic derivative $X(t) = g'(t).g(t)^{-1} \in \mathfrak{g}$. Differentiating $\bar{c}_1(t) = g(t).\bar{c}_2(t)$, we get

$$\bar{c}_1'(t) = g'(t).\bar{c}_2(t) + g(t).\bar{c}_2'(t),$$

and so

$$\bar{c}_1'(t) - g(t).\bar{c}_2'(t) = g'(t).\bar{c}_2(t) = X(t).g(t).\bar{c}_2(t) = X(t).\bar{c}_1(t).$$

Note that the left-hand side of this equation is orthogonal to the orbit through $\bar{c}_1(t)$, whereas the right-hand side is tangential to it (remember $T_{\bar{c}_1(t)}(G.\bar{c}_1(t)) = T_e l^{\bar{c}_1(t)}.\mathfrak{g}$), so both sides have to be zero. That means that $X(t)$ lies in the isotropy Lie algebra $\mathfrak{g}_{\bar{c}_1(t)}$ for each $t \in I$, and hence, by the result in the forgoing paragraph, $X(t)$ lies in the Lie algebra $\mathfrak{h}$ of H for all $t \in I$. But then $g(t)$ lies in a right coset of H for all $t \in I$. Obviously, this coset must be the same, say Hg, for all t. Consequently, we find $\bar{c}_1(t) = g.\bar{c}_2(t)$ for all t. □

CHAPTER 11

Lifting under weaker differentiability conditions

So far we have considered the lifting problem for either real analytic or smooth curves c in the orbit space $V/G = \sigma(V)$. In the smooth case we saw that one has to impose certain genericity conditions on c, see definition 10.1.6, in order to obtain a smooth lift to V. Now we want to tackle the problem under more general differentiability conditions for c. Otherwise put, let us forget about the mentioned genericity conditions and let us observe what we still can achieve. Note that, by 4.1.4 and 4.1.5, in generality, for a nongeneric curve c, there is no hope to get more than a twice differentiable lift $\bar{c}$.

This chapter presents the content of [**KLMR05**].

11.1. Lifting curves continuously

Let us remind of our setting: Let G be a compact Lie group and let $\rho : G \to \mathrm{O}(V)$ be an orthogonal representation in a real finite dimensional Euclidean vector space V with inner product $\langle . \mid . \rangle$. Let $\sigma_1, \ldots, \sigma_n$ be a system of homogeneous generators of $\mathbb{R}[V]^G$ with positive degrees $d_1, \ldots, d_n$, and consider the orbit map $\sigma = (\sigma_1, \ldots, \sigma_n) : V \to \mathbb{R}^n$. We may assume that $\sigma_1(v) = \langle v \mid v \rangle$.

In this section we shall lift curves continuously.

11.1.1. Theorem. *Let $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be continuous. Then there exists a global continuous lift $\bar{c} : \mathbb{R} \to V$ of c.*

Proof. We will use induction on the size of G. More precisely, recall that for two compact Lie groups G' and G we denote $G' < G$, if

- $\dim G' < \dim G$ or
- if $\dim G' = \dim G$, then G' has less connected components than G has.

In the simplest case, when $G = \{e\}$ is trivial, we find $\sigma(V) = V/G = V$, whence we can put $\bar{c} := c$.

Let us assume that for any $G' < G$ and any continuous $c : \mathbb{R} \to V/G'$ there exists a global continuous lift $\bar{c} : \mathbb{R} \to V$ of c, where $G' \to \mathrm{O}(V)$ is an orthogonal representation on an arbitrary real finite dimensional Euclidean vector space V.

We shall prove that then the same is true for G. Let $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be continuous. By lemma 10.1.2, we may remove the nontrivial fixed points of the G-action on V and suppose that $V^G = \{0\}$. The set $c^{-1}(0)$ is closed in $\mathbb{R}$ and, consequently, $c^{-1}(\sigma(V)\backslash\{0\}) = \mathbb{R}\backslash c^{-1}(0)$ is open in $\mathbb{R}$. Thus, we can write $c^{-1}(\sigma(V)\backslash\{0\}) = \bigcup_{i\in I}(a_i, b_i)$, where $a_i, b_i \in \mathbb{R} \cup \{\pm\infty\}$ with $a_i < b_i$ such that each (a_i, b_i) is maximal with respect to not containing zeros of c, and I is an at most countable set of indices. In particular, we have $c(a_i) = c(b_i) = 0$ for all $a_i, b_i \in \mathbb{R}$ appearing in the above presentation.

We assert that on each (a_i, b_i) there exists a continuous lift $\bar{c} : (a_i, b_i) \to V\backslash\{0\}$ of the restriction $c|_{(a_i,b_i)} : (a_i, b_i) \to \sigma(V)\backslash\{0\}$. In fact, since $V^G = \{0\}$, for all $v \in V\backslash\{0\}$ the isotropy groups G_v satisfy $G_v < G$. Therefore, by induction hypothesis and by theorem 9.5.3, we find local continuous lifts of $c|_{(a_i,b_i)}$ near any $t \in (a_i, b_i)$ and through all $v \in \sigma^{-1}(c(t))$. Suppose $\bar{c}_1 : (a_i, b_i) \supseteq (a, b) \to V\backslash\{0\}$

is a local continuous lift of $c|_{(a_i,b_i)}$ with maximal domain (a,b), where, say, $b<b_i$. Then, there exists a local continuous lift $\bar{c}_2$ of $c|_{(a_i,b_i)}$ near b, and there is a $t_0<b$ such that both $\bar{c}_1$ and $\bar{c}_2$ are defined near t_0. Since $\bar{c}_1(t_0)$ and $\bar{c}_2(t_0)$ are lying in the same orbit, there must exist a $g\in G$ such that $\bar{c}_1(t_0)=g.\bar{c}_2(t_0)$. But then,

$$\bar{c}_{12}(t):=\begin{cases} \bar{c}_1(t) & \text{for} \quad t\le t_0 \\ g.\bar{c}_2(t) & \text{for} \quad t\ge t_0 \end{cases}$$

is a local continuous lift of $c|_{(a_i,b_i)}$ defined on a larger interval than $\bar{c}_1$. Thus, we have shown that each local continuous lift of $c|_{(a_i,b_i)}$ defined on an open interval $(a,b)\subseteq(a_i,b_i)$ can be extended to a larger interval whenever $(a,b)\subsetneq(a_i,b_i)$. This proves the assertion.

Choosing a continuous lift of $c|_{(a_i,b_i)}$ for any $i\in I$ defines a continuous lift $\bar{c}:\mathbb{R}\backslash c^{-1}(0)\to V\backslash\{0\}$ of $c|_{\mathbb{R}\backslash c^{-1}(0)}$. We extend $\bar{c}$ to the whole of $\mathbb{R}$, by putting $\bar{c}(t):=0$ for $t\in c^{-1}(0)$. Note that, by $\sigma^{-1}(0)=\{0\}$, this is the only choice. The continuity of $\bar{c}$ at points $t\in c^{-1}(0)$ follows from the observation that $\sigma_1(\bar{c}(s))=\langle\bar{c}(s)\mid\bar{c}(s)\rangle=c_1(s)\to 0$ as $s\to t$. Therefore, $\bar{c}$ is a global continuous lift of c. This completes the induction and thus the proof. □

Remark. Note that proposition 2.3.2 yields a continuous lift $x=(x_1,\dots,x_n):\mathbb{R}\to\mathbb{R}^n$ of the continuous curve $P:\mathbb{R}\to\sigma(\mathbb{R}^n)=\mathbb{R}^n/\,\mathrm{S}_n$ which, moreover, lies in the fundamental domain

$$F=\{(y_1,\dots,y_n)\in\mathbb{R}^n: y_1\le y_2\le\dots\le y_n\}.$$

This is not possible in general, as the following example shows: Let $\mathbb{Z}_4$ act orthogonally on $\mathbb{R}^2$. Then every fundamental domain F is the interior of a right angle with apex at the origin, and by identifying the two sides we obtain the corresponding orbit space. If a curve in the orbit space crosses the line, where we have identified, then its continuous lift cannot lie entirely in F.

Nevertheless, it is easily seen that, if in the above theorem G is a finite reflection group, then we always can obtain a continuous lift $\bar{c}$ of c which is entirely contained in the closure of a fundamental domain.

11.1.2. Lifting paths. Theorem 11.1.1 is true also in a more general setting: Let X be a Hausdorff topological space, G a topological group, and $G\times X\to X$ a continuous action of G on X. Then we speak of a *G-space* X. If we replace the vector space V in theorem 11.1.1 by an arbitrary G-space X, where G is still a compact Lie group, then the conclusion is still valid.

The first step in proving this is the following lemma. This approach is due to Montgomery and Yang [**MY57**].

Lemma. *Suppose X is a G-space, G is a compact Lie group, and the orbit space X/G is homeomorphic to $I=[0,1]$. Then there is a* global cross section *for the projection $\pi:X\to X/G$, i.e., there exists a continuous map $\tau:X/G\to X$ with $\pi\circ\tau=\mathrm{id}_{X/G}$.*

Proof. By the remark below, X is compact and thus completely regular. One can show that there are slices at every point of a completely regular G-space, where G is a compact Lie group, see [**Bre72**, II.5.4].

It suffices to prove that π has a local cross section near each point of X/G. For, if $\tau_i:[\frac{i}{n},\frac{i+1}{n}]\to X$ is a cross section for $i=0,1,\dots,n-1$ and if $g_i\in G$ are such that $g_0=e$ and $g_i.\tau_i(\frac{i}{n})=\tau_{i-1}(\frac{i}{n})$ for $i=1,\dots,n-1$, then the map $\tau:I\to X$ with $\tau(t):=g_0g_1\cdots g_i.\tau_i(t)$ for $\frac{i}{n}\le t\le\frac{i+1}{n}$ is a global cross section. Similarly, if $J\subseteq I$ is an open subset, and if local cross sections exist near all points of J, then a cross section over J exists.

Now, using induction on the size of G, we can assume that the lemma is true for actions of any proper subgroup of G. Consider the space $F := X^G$ of fixed points under the G-action and its image $F^* := \pi(F) \subseteq I = X/G$ under π, which is closed in I since F is closed, see the lemma below. Now G acts on $X\backslash F$ without stationary points, i.e., points whose isotropy group is whole G, and with orbit space $I\backslash F^*$. Let $y \in X\backslash F$, $y^* := \pi(y)$, and let S be a slice at y. Since $G_y < G$ and $S/G_y \cong G.S/G$ is a neighborhood of y^*, the induction hypothesis, applied to the G_y-action on S, yields a local cross section at y for the orbit map $S \to S/G_y$ and hence for $X\backslash F \to I\backslash F^*$. As shown above, the existence of these local cross sections near all points of $X\backslash F$ implies the existence of a global cross section $\tau_0 : I\backslash F^* \to X\backslash F$ of the projection $X\backslash F \to I\backslash F^*$.

The image $C' := \tau_0(I\backslash F^*)$ is closed in $X\backslash F$: Let $(x_\alpha)_\alpha$ be a net in C' converging to $x \in X\backslash F$, then $x = \lim x_\alpha = \lim \tau_0(\pi(x_\alpha)) = \tau_0(\pi(x)) \in C'$. Consequently, $C := C' \cup F$ is closed in X. Clearly, C touches each orbit of X exactly once. So we can define the desired cross action $\tau : X/G \to X$ by $\{\tau(x^*)\} = G.x \cap C$ which is continuous, since for a closed $A \subseteq C$ also $\tau^{-1}(A) = \pi(A)$ is closed, by the lemma below. $\square$

Lemma. *Consider a G-space X, where G is a compact Lie group. Then the projection $\pi : X \to X/G$ is closed.*

Proof. Let $A \subseteq X$ be closed. Then $G.A$ is closed, since the action $l : G \times X \to X$ is closed: Let $C \subseteq G \times X$ be closed and let y be in the closure of $l(C)$, then there is a net (g_α, x_α) in C such that $l(g_\alpha, x_\alpha) = g_\alpha.x_\alpha$ converges to y. Passing to a subnet we may assume that g_α converges to g, since G is compact. Then, $x_\alpha = l(g_\alpha^{-1}, g_\alpha.x_\alpha)$ converges to $l(g^{-1}, y) = g^{-1}.y$. Thus, (g_α, x_α) converges to $(g, g^{-1}.y) \in C$, since C is closed. Thus, $y = l(g, g^{-1}.y) \in l(C)$.

But $G.A = \pi^{-1}(\pi(A))$, so $\pi(A)$ is closed. $\square$

Remark. The previous lemma implies that *for a G-space X, where G is a compact Lie group, the projection $\pi : X \to X/G$ is proper.* For: Let $K \subseteq X/G$ be compact and let $\{U_\alpha : \alpha \in A\}$ be an open covering of $\pi^{-1}(K)$. Since $\pi^{-1}(y)$ is compact for each $y \in X/G$, for each $y \in K$ there is a finite subset $A_y \subseteq A$ of indices such that the U_α, $\alpha \in A_y$, cover $\pi^{-1}(y)$. Put $U_y = \bigcup_{\alpha \in A_y} U_\alpha$ and $V_y = X/G\backslash\pi(X\backslash U_y)$. Since π is closed, V_y is open. Note that $\pi^{-1}(V_y) \subseteq U_y$, and $y \in V_y$. Let $V_{y_1}, \dots, V_{y_n}$ cover K. Then

$$\pi^{-1}(K) \subseteq \bigcup_{i=1}^n \pi^{-1}(V_{y_i}) \subseteq \bigcup_{i=1}^n U_{y_i} = \bigcup_{\substack{\alpha \in A_{y_i} \\ 1 \le i \le n}} U_\alpha.$$

Moreover, it is easy to verify that X/G is Hausdorff and that X/G is locally compact if and only if X is locally compact.

Now we can consider the general case.

Theorem. *Let X be a G-space, G a compact Lie group, and let $c : I \to X/G$ be a continuous curve. Then there exists a continuous lift $\bar{c} : I \to X$, i.e., $c = \pi \circ \bar{c}$.*

Proof. Consider $c^*X := X \times_{X/G} I$, the pullback of X via c:

$$\begin{tikzcd} c^*X \arrow[r, "c_1"] \arrow[d, "\pi_1"'] & X \arrow[d, "\pi"] \\ I \arrow[r, "c"'] & X/G \end{tikzcd}$$

G acts trivially on I, c_1 is the projection to X, π_1 is the projection to I, and c^*X is a G-space via $g.(x,t) := (g.x, g.t) = (g.x, t)$. Since π_1 is invariant, it induces a continuous map $\phi : (c^*X)/G \to I$. Now π_1 is open and onto, since π is, and thus ϕ is also open and onto. ϕ is injective: If (x,t) and (x',t) are both in c^*X, then $\pi(x) = c(t) = \pi(x')$, so that x and x' are in the same orbit, whence (x,t) and (x',t) are in the same orbit. Hence $\phi : (c^*X)/G \to I$ is a homeomorphism. Since ϕ is canonical, we may regard I as the orbit space $(c^*X)/G$. By the above lemma, there is a cross section $\tau : I \to c^*X$ and we have the following commutative diagram

$$\begin{tikzcd} c^*X \arrow[r, "c_1"] \arrow[d, shift left] & X \arrow[d, "\pi"] \\ I \arrow[u, shift left, "\tau"] \arrow[r, "c"'] & X/G \end{tikzcd}$$

Then, $\bar{c} := c_1 \circ \tau$ is a continuous lift of c. $\square$

11.2. The integer d

Before we continue the treatment of the lifting problem under finite differentiability conditions let us discuss the following number associated to any finite dimensional orthogonal representation of a compact Lie group. It will appear in the sequel when we formulate the regularity conditions for the curve in the orbit space.

11.2.1. Let $\rho : G \to O(V)$ be as in section 8.2. Choose a minimal system of homogeneous generators $\sigma_1, \dots, \sigma_n$ of positive degrees $d_1, \dots, d_n$ of $\mathbb{R}[V]^G$. We associate to ρ the following number:

$$d = d(\rho) := \max\{d_1, \dots, d_n\}.$$

The integer d is well-defined and independent of the choice of a minimal system of homogeneous generators of the algebra of invariant polynomials. This follows from the fact that a system of homogeneous invariants of positive degree generates $\mathbb{R}[V]^G$ as an algebra over $\mathbb{R}$ if and only if the images of the invariants in this system generate $\mathbb{R}[V]^G_+/(\mathbb{R}[V]^G_+)^2$ as a vector space over $\mathbb{R}$, where $\mathbb{R}[V]^G_+$ is the space of all invariants vanishing at the origin, see proof of lemma 9.6.6. The grading used here is given by the degree of the polynomials. Hence a system of homogeneous algebra generators has minimal cardinality if no generator is superfluous, and the number and the degrees of the elements in a minimal system of homogeneous generators are uniquely determined.

Note that the independence of d from the choice of a minimal system of homogeneous generators of $\mathbb{R}[V]^G$ also follows from the following lemma applied to the slice representation at 0.

11.2.2. Lemma. *Let $\rho : G \to O(V)$ be a finite dimensional representation of a compact Lie group G, let ρ' be some slice representation of ρ. Then, $d(\rho') \leq d(\rho)$.*

Proof. Let $\sigma_1, \dots, \sigma_n$ be a minimal system of homogeneous generators of $\mathbb{R}[V]^G$.

For an arbitrary $v \in V$ let $\rho' : G_v \to O(N_v)$ be its slice representation, and suppose S_v is a normal slice at v. Choose a minimal system of homogeneous generators $\tau_1, \dots, \tau_m$ of $\mathbb{R}[N_v]^{G_v}$ and assume that $\deg \tau_1 \leq \dots \leq \deg \tau_m = d(\rho')$. Then there exist polynomials $p_i \in \mathbb{R}[\mathbb{R}^m]$ such that

$$\sigma_i|_{S_v} = p_i(\tau_1|_{S_v}, \dots, \tau_m|_{S_v}) \qquad (1 \leq i \leq n).$$

On the other hand, by lemma 9.6.6, near $v \in N_v$ we have

$$\tau_j|_{S_v} = f_j(\sigma_1|_{S_v}, \dots, \sigma_n|_{S_v}) \qquad (1 \leq j \leq m),$$

where f_j are real analytic functions.

For contradiction assume $\deg \tau_m > d(\rho)$. Then all polynomials p_i do not depend on their last entry. Consequently, near $v \in N_v$,

$$\tau_m|_{S_v} = F(\tau_1|_{S_v}, \dots, \tau_{m-1}|_{S_v}),$$

where

$$F = f_m(p_1, \dots, p_n)$$

is real analytic. Introduce a new grading in $\mathbb{R}[\mathbb{R}^{m-1}]$ with respect to $\deg \tau_1 \le \cdots \le \deg \tau_{m-1}$ and write the function F as an infinite sum of homogeneous (with respect to this grading) terms. Let $\bar{F}$ be the sum of all terms of degree $\deg \tau_m$ in this presentation of F. We obtain, near $v \in N_v$,

$$\tau_m|_{S_v} = \bar{F}(\tau_1|_{S_v}, \dots, \tau_{m-1}|_{S_v}).$$

This means τ_m is a polynomial in $\tau_1, \dots, \tau_{m-1}$ in a neighborhood of v in N_v, and, hence, everywhere. This contradicts minimality of $\tau_1, \dots, \tau_m$. □

11.3. Lifting curves differentiably at each point

We are going to show that a sufficiently often differentiable curve in the orbit space allows local lifts near any $t_0 \in \mathbb{R}$ which are differentiable at t_0. It will be clarified soon what we mean by 'differentiable sufficiently often'.

Let $\sigma_1, \dots, \sigma_n$ be a minimal system of homogeneous generators of $\mathbb{R}[V]^G$. By lemma 10.1.2, we can assume without loss that $V^G = \{0\}$, and, thus, that $\sigma_1(v) = \langle v \mid v \rangle$.

11.3.1. Regular points. Again we start with the lifting problem at regular orbits. Taking advantage of the fact that $\pi : V_{\text{reg}} \to V_{\text{reg}}/G$ is a locally trivial fiber bundle, we can show, with exactly the same proof, the following variant of lemma 10.1.1.

Lemma. *A curve $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ of class C^d, where d is defined in section 11.2, admits an orthogonal lift $\bar{c}$ of class C^d in a neighborhood of a regular point $c(t_0) \in V_{\text{reg}}/G$. It is unique up to a transformation from G.* □

11.3.2. Multiplicity. We shall need the following stronger version of the multiplicity lemma 10.1.3:

Lemma. *Let $c = (c_1, \dots, c_n)$ be a curve in $\sigma(V) \subseteq \mathbb{R}^n$, where c_i is C^{d_i} with $d_i = \deg \sigma_i$, for $1 \le i \le n$, and $c(0) = 0$. Then the following two conditions are equivalent:*

(1) *$c_1(t) = t^2 c_{1,1}(t)$ near 0 for a continuous function $c_{1,1}$;*
(2) *$c_i(t) = t^{d_i} c_{i,i}(t)$ near 0 for a continuous function $c_{i,i}$, for all $1 \le i \le n$.*

Proof. The proof of the nontrivial implication (1) ⇒ (2) is the same as in the smooth case with $r = 1$, see the proof of lemma 10.1.3. The essential point is that the assumptions on the c_i to be in class C^{d_i} are just good enough to guarantee that $c_{(m)}(t) = (t^{-d_1 m} c_1(t), \dots, t^{-d_n m})$ is continuous. □

11.3.3. Proposition. *Let $c = (c_1, \dots, c_n) : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a curve of class C^d, where d is defined in section 11.2. Then, for any $t_0 \in \mathbb{R}$ there exists a local lift $\bar{c}$ of c near t_0 which is differentiable at t_0.*

Proof. We follow partially the algorithm given in the proof of theorem 10.1.5. Without loss of generality we may assume that $t_0 = 0$. We show the existence of local lifts of c which are differentiable at 0 through any $v \in \sigma^{-1}(c(0))$.

If $c(0) \neq 0$ corresponds to a regular orbit, then unique orthogonal C^d lifts exist through all $v \in \sigma^{-1}(c(0))$, by lemma 11.3.1.

If $c(0) = 0$, then c_1 must vanish of at least second order at 0, since $c_1(t) \geq 0$ for all $t \in \mathbb{R}$. That means $c_1(t) = t^2 c_{1,1}(t)$ near 0 for a continuous function $c_{1,1}$. By the variant of the multiplicity lemma 11.3.2, we find that $c_i(t) = t^{d_i} c_{i,i}(t)$ near 0 for $1 \leq i \leq n$, where $c_{1,1}, c_{2,2}, \dots, c_{n,n}$ are continuous functions. We consider the following continuous curve in $\sigma(V)$

$$\begin{aligned} c_{(1)}(t) &:= (c_{1,1}(t), c_{2,2}(t), \dots, c_{n,n}(t)) \\ &= (t^{-2} c_1(t), t^{-d_2} c_2(t), \dots, t^{-d_n} c_n(t)). \end{aligned}$$

By theorem 11.1.1, there exists a continuous lift $\bar{c}_{(1)}$ of $c_{(1)}$. Thus, $\bar{c}(t) := t \cdot \bar{c}_{(1)}(t)$ is a local lift of c near 0 which is differentiable at 0:

$$\sigma(\bar{c}(t)) = \sigma(t \cdot \bar{c}_{(1)}(t)) = (t^2 c_{1,1}(t), \dots, t^{d_n} c_{n,n}(t)) = c(t),$$

and

$$\lim_{t \to 0} \frac{t \cdot \bar{c}_{(1)}(t)}{t} = \lim_{t \to 0} \bar{c}_{(1)}(t) = \bar{c}_{(1)}(0).$$

Note that $\sigma^{-1}(0) = \{0\}$, therefore we are done in this case.

If $c(0) \neq 0$ corresponds to a singular orbit, let v be in $\sigma^{-1}(c(0))$ and consider the isotropy representation $G_v \to \mathrm{O}(N_v)$. By theorem 9.5.3, the lifting problem reduces to the same problem for C^d curves in N_v/G_v now passing through 0. By lemma 11.2.2, we can refer to the previous case, and the theorem is proved. □

11.4. Global differentiable lift

From the data of the previous section we shall construct a global differentiable lift to V of a C^d curve in the orbit space V/G. The number d was defined in section 11.2.

11.4.1. Lemma. *Consider a continuous curve $c : (a, b) \to X$ in a compact metric space X. Then the set A of all accumulation points of $c(t)$ as $t \searrow a$ is connected.*

Proof. For contradiction suppose that $A = A_1 \cup A_2$, where A_1 and A_2 are disjoint open and closed subsets of A. Since A is closed in X, also A_1 and A_2 are closed in X. There exist disjoint open subsets $A_1', A_2' \subseteq X$ with $A_1 \subseteq A_1'$ and $A_2 \subseteq A_2'$, because X is normal. Consider $F := X \backslash (A_1' \cup A_2')$ which is closed in X and hence compact. Since c visits A_1' and A_2' infinitely often and $c^{-1}(A_1')$ and $c^{-1}(A_2')$ are disjoint and open in $\mathbb{R}$, there has to exist a sequence $(t_m)_m \subseteq (a, b)$ with $t_m \to a$ and $c(t_m) \in F$ for all m. By the compactness of F, the sequence $(c(t_m))_m$ has an accumulation point y belonging to F. The point y is also an accumulation point of the curve $t \mapsto c(t)$ as $t \searrow a$. But this is a contradiction to $F \cap A = \emptyset$. □

11.4.2. Theorem. *Let $c = (c_1, \dots, c_n) : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a curve of class C^d, with d as in section 11.2. Then there exists a global differentiable lift $\bar{c} : \mathbb{R} \to V$ of c.*

Proof. The proof will be carried out by induction on the size of G.

If $G = \{e\}$ is trivial, then $\bar{c} := c$ is a global differentiable lift.

So let us assume that for any $G' < G$ and any $c : \mathbb{R} \to V/G'$ of class $C^{d'}$ there exists a global differentiable lift $\bar{c} : \mathbb{R} \to V$ of c, where $\rho' : G' \to \mathrm{O}(V)$ is an orthogonal representation on an arbitrary real finite dimensional Euclidean vector space V with $d' = d(\rho')$.

We shall prove that the same is true for G. Let $c = (c_1, \dots, c_n) : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be of class C^d. We may assume that $V^G = \{0\}$, by lemma 10.1.2. We can write $c^{-1}(\sigma(V) \backslash \{0\}) = \bigcup_{i \in J} (a_i, b_i)$, a disjoint union, where $a_i, b_i \in \mathbb{R} \cup \{\pm\infty\}$ with $a_i < b_i$ such that each (a_i, b_i) is maximal with respect to not containing zeros of c, and J is an at most countable set of indices. In particular, we have $c(a_i) = c(b_i) = 0$ for all $a_i, b_i \in \mathbb{R}$ appearing in the above presentation.

CLAIM. *On each (a_i, b_i) there exists a differentiable lift $\bar{c} : (a_i, b_i) \to V\backslash\{0\}$ of the restriction $c|_{(a_i,b_i)} : (a_i, b_i) \to \sigma(V)\backslash\{0\}$.*

The lack of nontrivial fixed points guarantees that for all $v \in V\backslash\{0\}$ the isotropy groups G_v satisfy $G_v < G$. Therefore, by induction hypothesis, theorem 9.5.3, and lemma 11.2.2, we find local differentiable lifts of $c|_{(a_i,b_i)}$ near any $t \in (a_i, b_i)$ and through all $v \in \sigma^{-1}(c(t))$. Suppose $\bar{c}_1 : (a_i, b_i) \supseteq (a, b) \to V\backslash\{0\}$ is a local differentiable lift of $c|_{(a_i,b_i)}$ with maximal domain (a, b), where, say, $b < b_i$. Then, there exists a local differentiable lift $\bar{c}_2$ of $c|_{(a_i,b_i)}$ near b, and there exists a $t_0 < b$ such that both $\bar{c}_1$ and $\bar{c}_2$ are defined near t_0. We may assume without loss that $\bar{c}_1(t_0) = \bar{c}_2(t_0) =: v_0$, by applying a transformation $g \in G$ to $\bar{c}_2$, say. We want to show that we can arrange the lift $\bar{c}_2$ in such a way that its derivative at t_0 matches with the derivative of $\bar{c}_1$ at t_0. We decompose

$$\bar{c}_i'(t_0) = \bar{c}_i'(t_0)^\top + \bar{c}_i'(t_0)^\perp \qquad i = 1, 2$$

into the parts tangent to the orbit $G.v_0$ and normal to it.

First we deal with the normal parts $\bar{c}_i'(t_0)^\perp \in V$. We consider the projection $p : G.S_{v_0} \cong G \times_{G_{v_0}} S_{v_0} \to G/G_{v_0} \cong G.v_0$ of a fiber bundle associated to the principal bundle $\pi : G \to G/G_{v_0}$, where S_{v_0} is a normal slice at v_0. Then, for t close to t_0, $\bar{c}_1$ and $\bar{c}_2$ are differentiable curves in $G.S_{v_0}$, whence $p \circ \bar{c}_i$ $(i = 1, 2)$ are differentiable curves in G/G_{v_0} which admit differentiable lifts g_i into G with $g_i(t_0) = e$ (via the horizontal lift of the principal connection, say). Consequently, $t \mapsto g_i(t)^{-1}.\bar{c}_i(t) =: \tilde{c}_i(t)$ are differentiable lifts of $c|_{(a_i,b_i)}$ near t_0 which lie in S_{v_0}, whence $\tilde{c}_i'(t_0) = \frac{d}{dt}\big|_{t=t_0} (g_i(t)^{-1}.\bar{c}_i(t)) = -g_i'(t_0).v_0 + \bar{c}_i'(t_0) \in N_{v_0}$. So, $\bar{c}_i'(t_0)^\top = (g_i'(t_0).v_0)^\top = g_i'(t_0).v_0$, and so for the normal part we get $\bar{c}_i'(t_0)^\perp = \tilde{c}_i'(t_0)$.

Since $\tilde{c}_i$ lie in S_{v_0} we can change to the isotropy representation $G_{v_0} \to O(N_{v_0})$ (using the same letters σ_i for the generators of $\mathbb{R}[N_{v_0}]^{G_{v_0}}$). We can suppose that $v_0 = 0$, i.e., $c(t_0) = 0$.

Let us remind of the continuous curve in $\sigma(V)$ defined in the proof of proposition 11.3.3 which depends on the point t_0:

$$c_{(1,t_0)}(t) := ((t - t_0)^{-2}c_1(t), (t - t_0)^{-d_2}c_2(t), \ldots, (t - t_0)^{-d_n}c_n(t)). \tag{11.1}$$

We find that for $i = 1, 2$:

$$\sigma(\tilde{c}_i'(t_0)) = \sigma\left(\lim_{t\to t_0} \frac{\tilde{c}_i(t) - \tilde{c}_i(t_0)}{t - t_0}\right) = \lim_{t\to t_0} \sigma\left(\frac{\tilde{c}_i(t)}{t - t_0}\right) = c_{(1,t_0)}(t_0),$$

i.e., $\tilde{c}_1'(t_0)$ and $\tilde{c}_2'(t_0)$ are lying in the same G_{v_0}-orbit. Thus, there must exist a $g_0 \in G_{v_0}$ such that $\bar{c}_1'(t_0)^\perp = \tilde{c}_1'(t_0) = g_0.\tilde{c}_2'(t_0) = g_0.\bar{c}_2'(t_0)^\perp = (g_0.\bar{c}_2)'(t_0)^\perp$.

Now we deal with the tangential parts. We search for a differentiable curve $t \mapsto g(t)$ in G with $g(t_0) = g_0$ and

$$\bar{c}_1'(t_0)^\top = \left(\tfrac{d}{dt}|_{t=t_0}(g(t).\bar{c}_2(t))\right)^\top = g'(t_0).v_0 + g_0.\bar{c}_2'(t_0)^\top.$$

But this linear equation can be solved for $g'(t_0)$, and, hence, the required curve $t \mapsto g(t)$ exists. Note that the normal parts still fit since

$$\left(\tfrac{d}{dt}|_{t=t_0}(g(t).\bar{c}_2(t))\right)^\perp = \left(g'(t_0).v_0 + g_0.\bar{c}_2'(t_0)\right)^\perp = 0 + g_0.\bar{c}_2'(t_0)^\perp = \bar{c}_1'(t_0)^\perp.$$

The two lifts $\bar{c}_1$ for $t \le t_0$ and $g.\bar{c}_2$ for $t \ge t_0$ fit together differentiably at t_0. This proves the claim.

Now let $\bar{c} : (a_i, b_i) \to V\backslash\{0\}$ be the differentiable lift of $c|_{(a_i,b_i)}$ constructed above. For $a_i \neq -\infty$, we put $\bar{c}(a_i) := 0$, the only choice. Consider the expression

$\gamma(t) := \frac{\bar{c}(t)}{t-a_i}$ which is a differentiable curve in $V\backslash\{0\}$ for $t \in (a_i, b_i)$. We want to show that $\lim_{t\searrow a_i}\gamma(t)$ exists. For t sufficiently close to a_i we have

$$\sigma(\gamma(t)) = \sigma\left(\frac{\bar{c}(t)}{t-a_i}\right) = c_{(1,a_i)}(t) \to c_{(1,a_i)}(a_i) \qquad \text{as } t \searrow a_i, \tag{11.2}$$

where now $c_{(1,a_i)}(t) := ((t-a_i)^{-2}c_1(t), (t-a_i)^{-d_2}c_2(t), \ldots, (t-a_i)^{-d_n}c_n(t))$. Let $\bar{c}_{(1,a_i)}$ be a continuous lift of $c_{(1,a_i)}$ which exists by theorem 11.1.1. Then (11.2) shows that the set A of all accumulation points of $(\gamma(t))_{t\searrow a_i}$ lies in the orbit $G.\bar{c}_{(1,a_i)}(a_i)$ through $\bar{c}_{(1,a_i)}(a_i)$. Lemma 11.4.1 gives that A is connected. In particular, the limit $\lim_{t\searrow a_i}\gamma(t)$ must exist, if G is a finite group. In general let us consider the projection $p : G.S_{v_1} \cong G \times_{G_{v_1}} S_{v_1} \to G/G_{v_1} \cong G.v_1$ of a fiber bundle associated to the principal bundle $\pi : G \to G/G_{v_1}$, where we choose $v_1 \in A$ and S_{v_1} is a normal slice at v_1. Then, for t close to a_i and $t > a_i$, the curve $t \mapsto \gamma(t)$ is differentiable in $G.S_{v_1}$, whence $t \mapsto p(\gamma(t))$ defines a differentiable curve in G/G_{v_1} which admits a differentiable lift $t \mapsto g(t)$ into G. Now, for t close to a_i, $t \mapsto g(t)^{-1}.\gamma(t)$ is a differentiable curve in S_{v_1} whose accumulation points for $t \searrow a_i$ have to lie in $G.v_1 \cap S_{v_1} = \{v_1\}$, since $\sigma(g(t)^{-1}.\gamma(t)) = \sigma(\gamma(t))$. That means that $t \mapsto g(t)^{-1}.\bar{c}(t)$ defines a differentiable lift of $c|_{(a_i,b_i)}$, for $t > a_i$ close to a_i, whose one-sided derivative at a_i exists:

$$\lim_{t\searrow a_i}\frac{g(t)^{-1}.\bar{c}(t)}{t-a_i} = \lim_{t\searrow a_i} g(t)^{-1}.\gamma(t) = v_1.$$

Let $t \mapsto g(t)$ be extended smoothly to (a_i, b_i) so that near b_i it is constant and replace $t \mapsto \bar{c}(t)$ by $t \mapsto g(t)^{-1}\bar{c}(t)$. Thus

$$\bar{c}'(a_i) := \lim_{t\searrow a_i}\frac{\bar{c}(t)}{t-a_i} = v_1.$$

The same reasoning is true for $b_i \neq +\infty$. Thus we have extended $\bar{c}$ differentiably to the closure of (a_i, b_i).

Let us finally construct a global differentiable lift of c defined on the whole of $\mathbb{R}$. For isolated points $t_0 \in c^{-1}(0)$ the two differentiable lifts on the neighboring intervals can be made to match differentiably, by applying a fixed transformation $g \in G$ to one of them, since the one-sided derivatives at t_0 both lie in the orbit through $\bar{c}_{(1,t_0)}(t_0)$ by a similar argument as in (11.2), where $\bar{c}_{(1,t_0)}$ is a continuous lift of the curve $c_{(1,t_0)}(t)$. Let E be the set of accumulation points of $c^{-1}(0)$. For connected components of $\mathbb{R}\backslash E$ we can proceed inductively to obtain differentiable lifts on them.

We extend the lift by 0 on the set E of accumulation points of $c^{-1}(0)$. Note that every lift $\tilde{c}$ of c has to vanish on E and is continuous there, since $\langle\tilde{c}(t) \mid \tilde{c}(t)\rangle = \sigma_1(\tilde{c}(t)) = c_1(t) \to 0$ as $t \to t'$ for $t' \in E$. We also claim that any lift $\tilde{c}$ of c is differentiable at any point $t' \in E$ with derivative 0. Namely, the difference quotient $t \mapsto \frac{\tilde{c}(t)}{t-t'}$ at t' is a lift of the curve $c_{(1,t')}$ which vanishes at t' by the following argument: Consider the local lift $\bar{c}$ of c near t' which is differentiable at t', provided by proposition 11.3.3. Let $(t_m)_{m\in\mathbb{N}} \subseteq c^{-1}(0)$ be a sequence with $t' \neq t_m \to t'$, consisting exclusively of zeros of c. Such a sequence always exists, since t' is an accumulation point of $c^{-1}(0)$. Then we have

$$\bar{c}'(t') = \lim_{t\to t'}\frac{\bar{c}(t)-\bar{c}(t')}{t-t'} = \lim_{m\to\infty}\frac{\bar{c}(t_m)}{t_m-t'} = 0.$$

Thus, we find $c_{(1,t')}(t') = \lim_{t\to t'}\sigma(\frac{\bar{c}(t)}{t-t'}) = \sigma(\bar{c}'(t')) = \sigma(0) = 0$.

This shows that extending our differentiable lift of c on $\mathbb{R}\backslash E$ by 0 at accumulation points of $c^{-1}(0)$ makes it a global differentiable lift on the whole of $\mathbb{R}$. So the induction and hence the proof is complete. □

11.4.3. Remark. The differentiability conditions of the curve c in the previous theorem 11.4.2 are best possible: In the case when the symmetric group S_n is acting in $\mathbb{R}^n$ by permuting the coordinates, and $\sigma_1, \dots, \sigma_n$ are the elementary symmetric polynomials with degrees $1, \dots, n$ (here $d = n$), there must not exist a differentiable lift, if the differentiability assumptions made on c are weakened, see the first example in 4.1.4.

CHAPTER 12

Twice differentiable lifts for polar representations

We show in this chapter that polar representations allow twice differentiable lifts to the representation space of sufficiently regular curves in the orbit space. This generalizes the result for the polynomial case presented in chapter 7. The crucial point there was Bronshtein's result (theorem 5.5.13; or Wakabayashi's version, theorem 6.3.1). We shall consider representations for which a generalization of Bronshtein's result holds true, and say that these representations have property ($\mathcal{B}$). It will turn out that polar representations automatically have property ($\mathcal{B}$).

This chapter is based on [**KLMR06**].

12.1. Property ($\mathcal{B}$)

12.1.1. Definition. We shall say that an orthogonal representation $\rho : G \to O(V)$ of a compact Lie group G on a real finite dimensional Euclidean vector space V has *property* ($\mathcal{B}_k$), if there exists a neighborhood $U = U(\rho)$ of 0 in $V/G = \sigma(V)$ such that each C^k curve in U admits a local differentiable lift $\bar{c}$ to V with locally bounded derivative.

Note that property ($\mathcal{B}_k$) is independent of the choice of generators of $\mathbb{R}[V]^G$. It is clear that, if a representation ρ has property ($\mathcal{B}_k$), then it has property ($\mathcal{B}_l$) for all $l \in \{k, k+1, \dots, \infty, \omega\}$ as well. We shall write simply property ($\mathcal{B}$), if the degree of differentiability k is not specified.

Example. The standard representation of the symmetric group S_n on $\mathbb{R}^n$ has property ($\mathcal{B}_n$). This follows from theorem 5.5.13 or from theorem 6.3.1.

12.1.2. Proposition. *Let $c = (c_1, \dots, c_n) : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a curve of class C^k in the orbit space of a representation $\rho : G \to O(V)$ with property ($\mathcal{B}_k$). Then for any $t_0 \in \mathbb{R}$ there exists a local differentiable lift $\bar{c}$ of c near t_0 with locally bounded derivative.*

Proof. For each $s \in \mathbb{R}\backslash\{0\}$ let us define a C^k curve $c_s : \mathbb{R} \to \sigma(V)$ by

$$c_s(t) = (s^{d_1} c_1(t), \dots, s^{d_n} c_n(t)).$$

There exists some $s = s(c; t_0) \in \mathbb{R}\backslash\{0\}$ such that $c_s(t) \in U$ for t near t_0, where U is the neighborhood of 0 in V/G introduced in the definition of property ($\mathcal{B}_k$). Since ρ has property ($\mathcal{B}_k$), there exists, near t_0, a local differentiable lift $\bar{c}_s$ of c_s to V with locally bounded derivative. Then, $\bar{c}(t) := s^{-1} \cdot \bar{c}_s(t)$ defines a local differentiable lift of c for t near t_0 whose derivative is locally bounded. $\square$

12.1.3. Proposition. *Assume that $\rho : G \to O(V)$ is a representation of a finite group G with property ($\mathcal{B}_k$). Then any slice representation ρ' of ρ has property ($\mathcal{B}_k$) as well.*

Proof. Let $\rho' : G_v \to O(N_v)$ be an arbitrary slice representation of ρ. Consider some normal slice S_v at v for the G-action on V. Then S_v/G_v is an open neighborhood of 0 in N_v/G_v which by theorem 9.4.4 is homeomorphic to $(G \times_{G_v} S_v)/G$ which in turn is an open neighborhood of $G.v$ in V/G.

Given a C^k curve c in S_v/G_v, we may view it as a curve in $(G\times_{G_v} S_v)/G$. Since ρ has property $(\mathcal{B}_k)$ and by proposition 12.1.2, there exists a local differentiable lift $\bar{c}$ of c to V with locally bounded derivative. The finiteness of G implies that $N_v = V$, and hence S_v is an open neighborhood of v in V. Therefore $\bar{c}$ is a local lift of c to N_v with respect to the G_v-action. □

12.1.4. The derivatives of lifts have unique norm.

LEMMA. *Let $c:\mathbb{R}\to V/G=\sigma(V)\subseteq\mathbb{R}^n$ be a curve in the orbit space V/G. We assume that G is finite. Let $t_0\in\mathbb{R}$. If $\bar{c}_1$ and $\bar{c}_2$ are lifts of c which are (one-sided) differentiable at t_0 and $\bar{c}_1(t_0)=\bar{c}_2(t_0)$, then there exists some $g\in G_{\bar{c}_1(t_0)}$ such that $\bar{c}_1'(t_0)=g.\bar{c}_2'(t_0)$.*

PROOF. Without loss we can assume that $t_0=0$.

Let $\bar{c}_1$ and $\bar{c}_2$ be lifts of $c:\mathbb{R}\to V/G$ which are (one-sided) differentiable at 0 and satisfy $\bar{c}_1(0)=\bar{c}_2(0)=: v_0$. We may suppose $V^G=\{0\}$, by lemma 10.1.2. We consider the following cases separately:

If $c(0)=0$, then $\bar{c}_1(0)=\bar{c}_2(0)=0$ and consequently, for $i=1,2$,

$$\sigma(\bar{c}_i'(0))=\sigma\left(\lim_{t\to 0}\frac{\bar{c}_i(t)}{t}\right)=\lim_{t\to 0}\sigma\left(\frac{\bar{c}_i(t)}{t}\right).$$

Now, for $t\neq 0$ we have $\sigma\left(\bar{c}_i(t)/t\right)=c_{(1)}(t)\in\sigma(V)$, where

$$c_{(1)}(t):=(t^{-d_1}c_1(t),\dots,t^{-d_n}c_n(t)).$$

Since $\sigma(V)$ is closed in $\mathbb{R}^n$ (see theorem 9.2.2), we find

$$\sigma(\bar{c}_i'(0))=\lim_{t\to 0}\sigma\left(\frac{\bar{c}_i(t)}{t}\right)=\lim_{t\to 0}c_{(1)}(t)\in\sigma(V),$$

i.e., σ maps $\bar{c}_1'(0)$ and $\bar{c}_2'(0)$ to the same point in $\sigma(V)$. (Note that, if only one-sided derivatives exist, then $t\to 0$ has to be replaced by $t\nearrow 0$ or $t\searrow 0$, respectively.) This shows that $\bar{c}_1'(0)$ and $\bar{c}_2'(0)$ lie in the same orbit, therefore we find some $g\in G=G_0$ with $\bar{c}_1'(0)=g.\bar{c}_2'(0)$.

If $c(0)\neq 0$: Since G is finite and therefore $N_{v_0}=V$, the ball S_{v_0} is a neighborhood of v_0 in V which contains the lifts $\bar{c}_1(t)$ and $\bar{c}_2(t)$ for t near 0. Hence, by theorem 9.5.3, we may change to the slice representation $G_{v_0}\to O(N_{v_0})$. Now we may assume that c is a curve in N_{v_0}/G_{v_0} with $c(0)=0$ and with lifts $\bar{c}_1(t)$ and $\bar{c}_2(t)$ to N_{v_0} for t near 0. So we refer to the former case. □

Note that lemma 12.1.4 does no longer hold, if finiteness of G is omitted:

EXAMPLE. Consider the standard action of SO(2) on $\mathbb{R}^2$. Then $\sigma(x_1,x_2)=x_1^2+x_2^2$ generates $\mathbb{R}[\mathbb{R}^2]^{\mathrm{SO}(2)}$ and $\mathbb{R}^2/\mathrm{SO}(2)=\sigma(\mathbb{R}^2)=[0,\infty)$. We consider the curve $c(t)=t^2$ and its differentiable lifts $\bar{c}_1(t)=(t,0)$ and $\bar{c}_2(t)=(t\cos t,t\sin t)$. We find $\bar{c}_1(2\pi)=\bar{c}_2(2\pi)=(2\pi,0)$, but $\bar{c}_1'(2\pi)=(1,0)$ and $\bar{c}_2'(2\pi)=(1,2\pi)$ cannot be transformed to each other by an element of $G_{(2\pi,0)}=\{\mathrm{id}\}$.

REMARK. If G is not finite, then the above lemma generalizes to the following statement: *Let $c:\mathbb{R}\to V/G=\sigma(V)\subseteq\mathbb{R}^n$ be a curve in the orbit space V/G. Let $t_0\in\mathbb{R}$. If $\bar{c}_1$ and $\bar{c}_2$ are lifts of c which are (one-sided) differentiable at t_0 and $\bar{c}_1(t_0)=\bar{c}_2(t_0)=: v_0$, then there exists some $g\in G_{v_0}$ such that $\bar{c}_1'(t_0)^\perp=g.\bar{c}_2'(t_0)^\perp$, where $\perp$ indicates the projection onto N_{v_0}.*

To see this: We consider the projection $p:G.S_{v_0}\cong G\times_{G_{v_0}}S_{v_0}\to G/G_{v_0}\cong G.v_0$ of a fiber bundle associated to the principal bundle $\pi:G\to G/G_{v_0}$, where S_{v_0} is a normal slice at v_0. Then, for t close to t_0, $\bar{c}_1$ and $\bar{c}_2$ are curves in $G.S_{v_0}$, whence $p\circ\bar{c}_i$ $(i=1,2)$ are curves in G/G_{v_0} which admit lifts g_i into G with $g_i(t_0)=e$, which are (one-sided) differentiable at t_0. Consequently, $t\mapsto g_i(t)^{-1}.\bar{c}_i(t)$ are lifts which lie in S_{v_0}, whence $\frac{d}{dt}\big|_{t=t_0}(g_i(t)^{-1}.\bar{c}_i(t)) = -g_i'(t_0).v_0 + \bar{c}_i'(t_0) \in N_{v_0}$. Thus, $\bar{c}_i'(t_0)^\perp = \frac{d}{dt}\big|_{t=t_0}(g_i(t)^{-1}.\bar{c}_i(t))$. By this observation, we may assume without loss that the lifts $\bar{c}_1$ and $\bar{c}_2$ lie in S_{v_0} for t close to t_0. Then the proof of lemma 12.1.4 gives the statement.

Remark. Lemma 12.1.4 implies that for any two differentiable lifts $\bar{c}_1$ and $\bar{c}_2$ of a curve c in V/G, where G is finite, we have $\|\bar{c}_1'(t)\| = \|\bar{c}_2'(t)\|$ for all t. So, if there exists some differentiable lift of c with locally bounded derivative, then any differentiable lift of c has this property as well.

12.1.5. Proposition. *Assume that $\rho : G \to O(V)$ is a representation of a finite group G with property $(\mathcal{B}_k)$. Let $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a curve of class C^k. Then there exists a global differentiable lift $\bar{c}$ of c to V with locally bounded derivative.*

Proof. Proposition 12.1.2 provides local differentiable lifts of c with locally bounded derivative near any $t \in \mathbb{R}$.

Now let us construct from these data a global differentiable lift of c with locally bounded derivative: First we glue the local differentiable lifts with locally bounded derivative just differentiably. It is sufficient to show that each local differentiable lift of c defined on an open interval I can be extended to a larger interval whenever $I \neq \mathbb{R}$.

Suppose that $\bar{c}_1 : I \to V$ is a local differentiable lift of c, and suppose the open interval I is bounded from above (say), and t_1 is its upper boundary point. Then, there exists a local differentiable lift $\bar{c}_2$ of c near t_1, and a $t_2 < t_1$ such that both $\bar{c}_1$ and $\bar{c}_2$ are defined near t_2. There is some $g \in G$ such that $\bar{c}_1(t_2) = g.\bar{c}_2(t_2)$. By lemma 12.1.4, we find an $h \in G_{\bar{c}_1(t_2)}$ with $\bar{c}_1'(t_2) = hg.\bar{c}_2'(t_2)$. Then $\bar{c}(t) := \bar{c}_1(t)$ for $t \le t_2$ and $\bar{c}(t) := hg.\bar{c}_2(t)$ for $t \ge t_2$ defines a differentiable lift of c on a larger interval.

Now let us show that the resulting global differentiable lift $\bar{c}$ of c has locally bounded derivative. Note first that by 12.1.4 the gluing process described above does not affect the local boundedness of the derivatives of the local lifts, provided by proposition 12.1.2. Let K be a compact subset of $\mathbb{R}$. The domains of definition of the local lifts constitute an open covering of K which contains a finite open subcovering $\{I_j\}$. By shrinking the open intervals I_j in the subcovering a bit we can assume that K is covered by finitely many compact intervals K_i each of which lies in some I_j. Since the local differentiable lifts have locally bounded derivatives, there exist constants C_{K_i} for all i such that

$$\|\bar{c}'(t)\| \le C_{K_i} \qquad \text{for all } t \in K_i.$$

If we put $C_K := \max\{C_{K_i}\}$, then

$$\|\bar{c}'(t)\| \le C_K \qquad \text{for all } t \in K.$$

This completes the proof. □

12.2. Stability of property ($\mathcal{B}$)

We shall prove that property $(\mathcal{B}_k)$ is stable under passing to subrepresentations and building orthogonal direct sums of representations.

12.2.1. Proposition. *Let $\rho : G \to O(V)$ be an orthogonal representation of a compact Lie group on a real finite dimensional Euclidean vector space V having property $(\mathcal{B}_k)$. For any G-invariant linear subspace $W \subseteq V$ the subrepresentation $\rho' : G \to O(W)$ has property $(\mathcal{B}_k)$ as well.*

PROOF. The restriction map $\mathbb{R}[V]^G \to \mathbb{R}[W]^G : p \mapsto p|_W$ is a surjective algebra homomorphism. Hence, if $\sigma_1, \dots, \sigma_n$ are generators of $\mathbb{R}[V]^G$, then their restrictions $\sigma_1|_W, \dots, \sigma_n|_W$ generate $\mathbb{R}[W]^G$. By lemma 8.2.2, the map $\sigma|_W = (\sigma_1|_W, \dots, \sigma_n|_W)$ induces a homeomorphism between the orbit space W/G and the image $\sigma|_W(W)$. Then the orbit space $W/G = \sigma|_W(W)$ is naturally a subset of the orbit space $V/G = \sigma(V)$.

Let $c : \mathbb{R} \to \sigma|_W(W) \cap U$ be a C^k curve in the orbit space $\sigma|_W(W)$, where $U = U(\rho)$ is the open neighborhood of 0 in $\sigma(V)$ from the definition of property $(\mathcal{B}_k)$ (see definition 12.1.1). We may view c as a curve in the orbit space $\sigma(V)$, and since the representation ρ has property $(\mathcal{B}_k)$, we can lift c to a local differentiable curve $\bar{c}$ in V with locally bounded derivative. But then $\bar{c}$ has obviously to lie in the G-invariant subspace W. This completes the proof. $\square$

12.2.2. PROPOSITION. *Suppose that $\rho_i : G_i \to O(V_i)$, for $1 \le i \le l$, are orthogonal representations of compact Lie groups G_i on real finite dimensional Euclidean vector spaces V_i having property $(\mathcal{B}_{k_i})$. Then the orthogonal direct sum*

$$\rho_1 \oplus \dots \oplus \rho_l : G_1 \times \dots \times G_l \longrightarrow O(V_1 \oplus \dots \oplus V_l)$$

of the representations $\rho_1, \dots, \rho_l$ has property $(\mathcal{B}_k)$, where $k = \max\{k_1, \dots, k_l\}$.

PROOF. It is sufficient to consider the case $l = 2$, since the general case follows by induction.

If $\langle \ \mid \ \rangle_1$ and $\langle \ \mid \ \rangle_2$ denote the inner products on V_1 and V_2, then

$$\langle v_1 + v_2 \mid w_1 + w_2 \rangle := \langle v_1 \mid w_1 \rangle_1 + \langle v_2 \mid w_2 \rangle_2$$

defines an inner product on $V = V_1 \oplus V_2$ which makes V_1 and V_2 into orthogonal subspaces of V. The action of $G = G_1 \times G_2$ on $V_1 \oplus V_2$ is obviously again orthogonal. Moreover, we find $\mathbb{R}[V]^G = \mathbb{R}[V_1 \oplus V_2]^{G_1 \times G_2} \cong \mathbb{R}[V_1]^{G_1} \otimes \mathbb{R}[V_2]^{G_2}$ and $V/G = (V_1 \oplus V_2)/(G_1 \times G_2) \cong V_1/G_1 \times V_2/G_2$. Now any C^k curve c in $U_1 \times U_2 \subseteq V/G$ has the form $c = (c_1, c_2)$ for C^k curves c_i in $U_i \subseteq V_i/G_i$, where $U_i = U(\rho_i)$ as in definition 12.1.1, which allow local differentiable lifts $\bar{c}_i$ with locally bounded derivative to V_i, by assumption. This shows that $\rho = \rho_1 \oplus \rho_2$ has property $(\mathcal{B}_k)$. $\square$

12.3. C^1 lifts

12.3.1. Multiplicity. We state another variant of the multiplicity lemmas 10.1.3 and 11.3.2 with essentially the same proof.

Let $d = d(\rho)$ be the integer associated to $\rho : G \to \mathrm{O}(V)$ in section 11.2.

LEMMA. *Let $c = (c_1, \dots, c_n)$ be a curve in $\sigma(V) \subseteq \mathbb{R}^n$ of class C^r, where $r \ge d$, and $c(0) = 0$. Then the following two conditions are equivalent:*

(1) *$c_1(t) = t^2 c_{1,1}(t)$ near 0 for a C^{r-2} function $c_{1,1}$;*

(2) *$c_i(t) = t^{d_i} c_{i,i}(t)$ near 0 for a C^{r-d_i} function $c_{i,i}$, for all $1 \le i \le n$.* $\square$

12.3.2. PROPOSITION. *Assume that $\rho : G \to \mathrm{O}(V)$ is a representation of a finite group G with property $(\mathcal{B}_k)$. Let $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a curve of class C^{k+d}. Then for any $t_0 \in \mathbb{R}$ there exists a local differentiable lift $\bar{c}$ of c near t_0 whose derivative is continuous at t_0.*

PROOF. Without loss of generality we may assume that $t_0 = 0$. We show the existence of local differentiable lifts of c whose derivatives are continuous at 0 through any $v \in \sigma^{-1}(c(0))$. By lemma 10.1.2 we can assume $V^G = \{0\}$.

If $c(0) \neq 0$ corresponds to a regular orbit, then unique orthogonal C^{k+d} lifts defined near 0 exist through all $v \in \sigma^{-1}(c(0))$, by lemma 11.3.1.

If $c(0) = 0$, then c_1 must vanish of at least second order at 0, since $c_1(t) \ge 0$ for all $t \in \mathbb{R}$. That means $c_1(t) = t^2 c_{1,1}(t)$ near 0 for a C^{k+d-2} function $c_{1,1}$. By the multiplicity lemma 12.3.1 we find that $c_i(t) = t^{d_i} c_{i,i}(t)$ near 0 for $1 \le i \le n$, where $c_{1,1}, c_{2,2}, \ldots, c_{n,n}$ are functions of class $C^{k+d-2}, C^{k+d-d_2}, \ldots, C^{k+d-d_n}$, respectively. We consider the following C^k curve in $\sigma(V)$ (since $\sigma(V)$ is closed in $\mathbb{R}^n$, see theorem 9.2.2):

$$\begin{aligned} c_{(1)}(t) &:= (c_{1,1}(t), c_{2,2}(t), \ldots, c_{n,n}(t)) \\ &= (t^{-2} c_1(t), t^{-d_2} c_2(t), \ldots, t^{-d_n} c_n(t)). \end{aligned}$$

By property $(\mathcal{B}_k)$ and proposition 12.1.2, there exists a local differentiable lift $\bar{c}_{(1)}$ of $c_{(1)}$ with locally bounded derivative. Thus, $\bar{c}(t) := t \cdot \bar{c}_{(1)}(t)$ is a local differentiable lift of c near 0 with derivative $\bar{c}'(t) = \bar{c}_{(1)}(t) + t\bar{c}'_{(1)}(t)$ which is continuous at $t = 0$ with $\bar{c}'(0) = \bar{c}_{(1)}(0)$. Note that $\sigma^{-1}(0) = \{0\}$, therefore we are done in this case.

If $c(0) \neq 0$ corresponds to a singular orbit, let v be in $\sigma^{-1}(c(0))$ and consider the slice representation $G_v \to \mathrm{O}(N_v)$. By theorem 9.5.3, the lifting problem reduces to the same problem for curves in N_v/G_v now passing through 0. By proposition 12.1.3 we may refer to the former case. □

12.3.3. THEOREM. *Assume that $\rho : G \to \mathrm{O}(V)$ is a representation of a finite group G with property $(\mathcal{B}_k)$. Let $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a curve of class C^{k+d}. Then any differentiable lift $\bar{c}$ of c is actually of class C^1.*

PROOF. Let $\bar{c}$ be a differentiable lift of c. Let $t_0 \in \mathbb{R}$ be arbitrary. We show that $\bar{c}'$ is continuous at t_0. Let $\tilde{c}$ denote the local differentiable lift of c near t_0 with continuous derivative at t_0, provided by proposition 12.3.2. Consider a sequence $(t_m)_m \subseteq \mathbb{R}$ with $t_m \to t_0$. For every m there is a $g_m \in G$ such that $\bar{c}(t_m) = g_m.\tilde{c}(t_m)$. Since G is finite, we may choose a subsequence of $(t_m)_m$ again denoted by $(t_m)_m$ such that $\bar{c}(t_m) = g.\tilde{c}(t_m)$ for some fixed $g \in G$ and all m. By lemma 12.1.4, there exist $h_m \in G_{\bar{c}(t_m)}$ with $\bar{c}'(t_m) = h_m g.\tilde{c}'(t_m)$ for all m. Passing again to a subsequence we find a fixed $h \in G_{\bar{c}(t_m)}$ such that $\bar{c}(t_m) = h.\bar{c}(t_m) = hg.\tilde{c}(t_m)$ and $\bar{c}'(t_m) = hg.\tilde{c}'(t_m)$ for all m. Then

$$\bar{c}(t_0) = \lim_{t_m \to t_0} \bar{c}(t_m) = \lim_{t_m \to t_0} hg.\tilde{c}(t_m) = hg. \lim_{t_m \to t_0} \tilde{c}(t_m) = hg.\tilde{c}(t_0)$$

and

$$\bar{c}'(t_0) = \lim_{t_m \to t_0} \frac{\bar{c}(t_m) - \bar{c}(t_0)}{t_m - t_0} = \lim_{t_m \to t_0} \frac{hg.\tilde{c}(t_m) - hg.\tilde{c}(t_0)}{t_m - t_0} = hg.\tilde{c}'(t_0)$$

and hence

$$\lim_{t_m \to t_0} \bar{c}'(t_m) = \lim_{t_m \to t_0} hg.\tilde{c}'(t_m) = hg.\tilde{c}'(t_0) = \bar{c}'(t_0).$$

This completes the proof. □

The forgoing theorem 12.3.3 is false, if G is not finite:

EXAMPLE. Again consider the standard action of $\mathrm{SO}(2)$ on $\mathbb{R}^2$ with orbit map $\sigma(x_1, x_2) = x_1^2 + x_2^2$. Let us consider the curve $c(t) = t^4$ and its differentiable lift $\bar{c}(t) = \big(t^2 \cos \frac{1}{t}, t^2 \sin \frac{1}{t}\big)$. But the derivative $\bar{c}'(t) = \big(2t \cos \frac{1}{t} + \sin \frac{1}{t}, 2t \sin \frac{1}{t} - \cos \frac{1}{t}\big)$ is not continuous at $t = 0$.

REMARK. The failure of theorem 12.3.3 in this special example really is due to the fact that $\mathrm{SO}(2)$ is infinite, since there is the following result due to Bony [**Bon05**]: *Any non-negative function $f : \mathbb{R} \to \mathbb{R}$ of class C^{2m} can be represented as sum of squares $f = g^2 + h^2$ of C^m functions g and h.* This result implies that the standard representation of $\mathrm{SO}(2)$ on $\mathbb{R}^2$ has property $(\mathcal{B}_2)$, and hence any standard representation of $\mathrm{SO}(n)$ on $\mathbb{R}^n$ $(n \geq 2)$ has property $(\mathcal{B}_2)$ as well. But see example 12.5.2.

Note that the lifting problem for the standard representation of $\mathrm{SO}(n)$ on $\mathbb{R}^n$ is just the problem of representing non-negative functions defined on $\mathbb{R}$ as sums of squares. This question is related to *Hilbert's 17th problem.* Whether it is possible, to write a positive semidefinite rational function in indeterminates over the reals, as a sum of squares of rational functions in indeterminates over the reals? The answer is yes, and it was proved by Emil Artin in 1927. Additionally, Artin showed that the answer is also yes if the reals were replaced by the rationals. In [**FP78**], while proving their celebrated inequality, Fefferman and Phong stated and sketchily proved that any non-negative $C^{3,1}$ function in $\mathbb{R}^n$ is a sum of squares of $C^{1,1}$ functions. For $n \geq 4$ this result is sharp: In [**BBCP06**] is shown that there exist non-negative smooth functions in $\mathbb{R}^n$ that are not sums of squares of C^2 functions. The core of the proof is a result due to Hilbert [**Hil88**] asserting that there are homogeneous polynomials of degree 4 that are not sums of squares of polynomials. For analogous reasons there exist smooth non-negative functions in $\mathbb{R}^3$ that are not sums of squares of C^3 functions. In dimensions 1 and 2 there are no algebraic obstacles to the decomposition into sums of squares. In dimension 2 any non-negative C^4 function f is a sum of squares of C^2 functions, if $f(t) = \nabla^2 f(t) = 0$ implies $\nabla^4 f(t) = 0$ ([**Bon05**]). In the result due to Bony mentioned in the previous paragraph which deals with the one-dimensional case the regularity of g and h is best possible. If $f : \mathbb{R} \to \mathbb{R}$ is non-negative and smooth, this result allows to write $f = g_m^2 + f_m^2$ with C^m functions g_m and h_m for all m. However, this decomposition depends on m and it is not clear whether f can be represented as sum of squares of smooth functions. In [**BCR87**] and [**Bru79**] counter-examples due to Cohen and Epstein are mentioned, but it seems that they have never been published.

12.4. Twice differentiable lifts

12.4.1. PROPOSITION. *Assume that $\rho : G \to \mathrm{O}(V)$ is a representation of a finite group G with property $(\mathcal{B}_k)$. Let $c = (c_1, \dots, c_n) : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a curve of class C^{k+2d}. Then for any $t_0 \in \mathbb{R}$ there exists a local C^1 lift $\bar{c}$ of c near t_0 which is twice differentiable at t_0.*

PROOF. Without loss of generality we may assume that $t_0 = 0$. We show the existence of local C^1 lifts of c which are twice differentiable at 0 through any $v \in \sigma^{-1}(c(0))$. By lemma 10.1.2 we can assume $V^G = \{0\}$.

If $c(0) \neq 0$ corresponds to a regular orbit, then unique orthogonal C^{k+2d} lifts defined near 0 exist through all $v \in \sigma^{-1}(c(0))$, by lemma 11.3.1.

If $c(0) = 0$, then as in the proof of proposition 12.3.2 we find that the curve

$$\begin{aligned} c_{(1)}(t) &:= (c_{1,1}(t), c_{2,2}(t), \dots, c_{n,n}(t)) \\ &= (t^{-2}c_1(t), t^{-d_2}c_2(t), \dots, t^{-d_n}c_n(t)) \end{aligned}$$

lies in $\sigma(V)$ and is of class C^{k+d}. By property $(\mathcal{B}_k)$ and theorem 12.3.3, there exists a local C^1 lift $\bar{c}_{(1)}$ of $c_{(1)}$. Thus, $\bar{c}(t) := t \cdot \bar{c}_{(1)}(t)$ is a local C^1 lift of c near 0 with derivative $\bar{c}'(t) = \bar{c}_{(1)}(t) + t\bar{c}'_{(1)}(t)$ which is differentiable at $t = 0$:

$$\lim_{t\to 0} \frac{\bar{c}'(t) - \bar{c}'(0)}{t} = \lim_{t\to 0} \frac{\bar{c}_{(1)}(t) - \bar{c}_{(1)}(0) + t\bar{c}'_{(1)}(t)}{t} = 2\bar{c}'_{(1)}(0).$$

If $c(0) \neq 0$ corresponds to a singular orbit, let v be in $\sigma^{-1}(c(0))$ and consider the isotropy representation $G_v \to \mathrm{O}(N_v)$. By theorem 9.5.3, the lifting problem reduces to the same problem for curves in N_v/G_v now passing through 0. By proposition 12.1.3 we may refer to the former case. □

12.4.2. THEOREM. *Assume that $\rho : G \to \mathrm{O}(V)$ is a representation of a finite group G with property $(\mathcal{B}_k)$. Let $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a curve of class C^{k+2d}. Then there exists a global twice differentiable lift $\bar{c}$ of c.*

PROOF. The proof will be carried out by induction on the cardinality of G.

If $G = \{e\}$ is trivial, then $\bar{c} := c$ is a global twice differentiable lift.

So let us assume that for any finite G' with $|G'| < |G|$ and any $c : \mathbb{R} \to V/G'$ of class $C^{k+2d'}$ there exists a global twice differentiable lift $\bar{c} : \mathbb{R} \to V$ of c, where $\rho' : G' \to \mathrm{O}(V)$ is an orthogonal representation on an arbitrary real finite dimensional Euclidean vector space V with property $(\mathcal{B}_k)$, and $d' = d(\rho')$.

We shall prove that the same is true for G. Let $c = (c_1, \ldots, c_n) : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be of class C^{k+2d}. We may assume that $V^G = \{0\}$, by lemma 10.1.2. We can write $c^{-1}(\sigma(V)\backslash\{0\}) = \bigcup_i (a_i, b_i)$, a disjoint, at most countable union, where $a_i, b_i \in \mathbb{R} \cup \{\pm\infty\}$ with $a_i < b_i$ such that each (a_i, b_i) is maximal with respect to not containing zeros of c. In particular, we have $c(a_i) = c(b_i) = 0$ for all $a_i, b_i \in \mathbb{R}$ appearing in the above presentation.

CLAIM. *On each (a_i, b_i) there exists a twice differentiable lift $\bar{c} : (a_i, b_i) \to V\backslash\{0\}$ of the restriction $c|_{(a_i,b_i)} : (a_i, b_i) \to \sigma(V)\backslash\{0\}$*

The lack of nontrivial fixed points guarantees that for all $v \in V\backslash\{0\}$ the isotropy groups G_v satisfy $|G_v| < |G|$. Therefore, by induction hypothesis, which is fulfilled by proposition 12.1.3 and lemma 11.2.2, and by theorem 9.5.3, we find local twice differentiable lifts of $c|_{(a_i,b_i)}$ near any $t \in (a_i, b_i)$ and through all $v \in \sigma^{-1}(c(t))$. Suppose that $\bar{c}_1 : (a_i, b_i) \supseteq (a, b) \to V\backslash\{0\}$ is a local twice differentiable lift of $c|_{(a_i,b_i)}$ with maximal domain (a, b), where, say, $b < b_i$. Then there exists a local twice differentiable lift $\bar{c}_2$ of $c|_{(a_i,b_i)}$ near b, and there exists a $t_0 < b$ such that both $\bar{c}_1$ and $\bar{c}_2$ are defined near t_0. Let $(t_m)_m$ be a sequence with $t_m \to t_0$. For any m there exists a $g_m \in G$ such that $\bar{c}_1(t_m) = g_m.\bar{c}_2(t_m)$. Since G is finite, we may choose a subsequence again denoted by $(t_m)_m$ such that $\bar{c}_1(t_m) = g.\bar{c}_2(t_m)$ for a fixed $g \in G$ and for all m. By lemma 12.1.4, there are $h_m \in G_{\bar{c}_1(t_m)}$ with $\bar{c}_1'(t_m) = h_m g.\bar{c}_2'(t_m)$ for all m. Passing again to a subsequence we find a fixed $h \in G_{\bar{c}_1(t_m)}$ such that $\bar{c}_1(t_m) = g.\bar{c}_2(t_m) = hg.\bar{c}_2(t_m)$ and $\bar{c}_1'(t_m) = hg.\bar{c}_2'(t_m)$ for all m. Consequently,

$$\bar{c}_1(t_0) = \lim_{t_m \to t_0} \bar{c}_1(t_m) = \lim_{t_m \to t_0} hg.\bar{c}_2(t_m) = hg.\bar{c}_2(t_0)$$

and

$$\bar{c}_1'(t_0) = \lim_{t_m \to t_0} \bar{c}_1'(t_m) = \lim_{t_m \to t_0} hg.\bar{c}_2'(t_m) = hg.\bar{c}_2'(t_0),$$

and hence

$$\bar{c}_1''(t_0) = \lim_{t_m \to t_0} \frac{\bar{c}_1'(t_m) - \bar{c}_1'(t_0)}{t_m - t_0} = \lim_{t_m \to t_0} \frac{hg.\bar{c}_2'(t_m) - hg.\bar{c}_2'(t_0)}{t_m - t_0} = hg.\bar{c}_2''(t_0).$$

So $\bar{c}(t) := \bar{c}_1(t)$ for $t \le t_0$ and $\bar{c}(t) := hg.\bar{c}_2(t)$ for $t \ge t_0$ defines a twice differentiable lift of $c|_{(a_i,b_i)}$ on a larger interval than (a, b). This proves the claim.

Now let $\bar{c} : (a_i, b_i) \to V\backslash\{0\}$ be the twice differentiable lift of $c|_{(a_i,b_i)}$ constructed above. For $a_i \ne -\infty$, we put $\bar{c}(a_i) := 0$ and $\bar{c}'(a_i) := \lim_{t \searrow a_i} \frac{\bar{c}(t)}{t - a_i}$ which exists as shown in the proof of theorem 11.4.2. Then $\bar{c}$ is one-sided continuous at a_i, since $\langle \bar{c}(t) \mid \bar{c}(t) \rangle = \sigma_1(\bar{c}(t)) = c_1(t)$. Let $\tilde{c}$ be a local C^1 lift of c defined near a_i which is twice differentiable at a_i, provided by proposition 12.4.1. Then we find

$$\lim_{t \searrow a_i} \bar{c}(t) = \bar{c}(a_i) = 0 = \tilde{c}(a_i).$$

Let $(t_m)_m \subseteq (a_i, b_i)$ be a sequence with $t_m \searrow a_i$. By the arguments above, we may pass to a subsequence and find a $g \in G$ and an $h \in G_{\bar{c}(t_m)}$ such that $\bar{c}(t_m) = g.\tilde{c}(t_m) = hg.\tilde{c}(t_m)$ and $\bar{c}'(t_m) = hg.\tilde{c}'(t_m)$ for all m. Therefore we have

$$\bar{c}'(a_i) = \lim_{t_m \searrow a_i} \frac{\bar{c}(t_m)}{t_m - a_i} = \lim_{t_m \searrow a_i} \frac{hg.\tilde{c}(t_m)}{t_m - a_i} = hg.\tilde{c}'(a_i). \tag{12.1}$$

Moreover,

$$\lim_{t_m \searrow a_i} \bar{c}'(t_m) = \lim_{t_m \searrow a_i} hg.\tilde{c}'(t_m) = hg.\tilde{c}'(a_i) = \bar{c}'(a_i),$$

since $\tilde{c}$ is C^1. It follows that the set of all accumulation points of $(\bar{c}'(t))_{t \searrow a_i}$ lies in the orbit $G.\tilde{c}'(a_i)$. Since G is finite, lemma 11.4.1 implies that $\bar{c}'(t)$ converges for $t \searrow a_i$, with limit $\bar{c}'(a_i)$, because it does so along the sequence $(t_m)_m$. Otherwise put, the lift $\bar{c}$ is continuously differentiable also at the boundary point a_i of its domain.

For the sequence $(t_m)_m$ from above we can argue further

$$\frac{\bar{c}'(t_m) - \bar{c}'(a_i)}{t_m - a_i} = \frac{hg.\tilde{c}'(t_m) - hg.\tilde{c}'(a_i)}{t_m - a_i} \to hg.\tilde{c}''(a_i) \qquad \text{as } t_m \searrow a_i,$$

since the lift $\tilde{c}$ is twice differentiable at a_i. Hence the set of all accumulation points of $\left(\frac{\bar{c}'(t)-\bar{c}'(a_i)}{t-a_i}\right)_{t\searrow a_i}$ is a subset of $G_{\bar{c}'(a_i)}hg.\tilde{c}''(a_i)$: Any accumulation point of $\left(\frac{\bar{c}'(t)-\bar{c}'(a_i)}{t-a_i}\right)_{t\searrow a_i}$ corresponds to a sequence $(t_m)_m \in (a_i, b_i)$ with $t_m \searrow a_i$ such that $\frac{\bar{c}'(t_m)-\bar{c}'(a_i)}{t_m-a_i} \to \hat{h}\hat{g}.\tilde{c}''(a_i)$, where $\hat{h}$ and $\hat{g}$ are found by repeating the procedure above. From the equation $\hat{h}\hat{g}.\tilde{c}'(a_i) = \bar{c}'(a_i) = hg.\tilde{c}'(a_i)$, which follows from (12.1), we can read off $(hg)^{-1}\hat{h}\hat{g} \in G_{\tilde{c}'(a_i)} = (hg)^{-1}G_{\bar{c}'(a_i)}hg$, and hence $\hat{h}\hat{g} \in G_{\bar{c}'(a_i)}hg$.

By lemma 11.4.1, we have that $\frac{\bar{c}'(t)-\bar{c}'(a_i)}{t-a_i}$ converges for $t \searrow a_i$, with limit $hg.\tilde{c}''(a_i)$, since it does so along the sequence $(t_m)_m$. That means that the one-sided second derivative of $\bar{c}$ exists at a_i. The same reasoning is true for $b_i \neq +\infty$. So we have extended our lift $\bar{c}$ twice differentiably to the closure of (a_i, b_i).

Let us now construct a global twice differentiable lift of c defined on the whole of $\mathbb{R}$. For isolated points $t_0 \in c^{-1}(0)$ the two twice differentiable lifts on the neighboring intervals can be made to match twice differentiably, by applying a fixed transformation from G to one of them: Let $\bar{c}_1$ and $\bar{c}_2$ denote the lifts left and right of t_0. Then $\bar{c}_1(t_0) = \bar{c}_2(t_0) = 0$ and, by lemma 12.1.4, we find some $g \in G$ such that $\bar{c}_1'(t_0) = g.\bar{c}_2'(t_0)$. Let $\tilde{c}$ be the local C^1 lift near t_0 which is twice differentiable at t_0, provided by proposition 12.4.1. By the same argumentation as in the previous paragraph we find $h_1, h_2 \in G$ such that

$$h_1.\tilde{c}'(t_0) = \bar{c}_1'(t_0) = g.\bar{c}_2'(t_0) = h_2.\tilde{c}'(t_0),$$

and for the one-sided second derivatives we have

$$\lim_{t \nearrow t_0} \frac{\bar{c}_1'(t) - \bar{c}_1'(t_0)}{t - t_0} = h_1.\tilde{c}''(t_0) \quad \text{and} \quad \lim_{t \searrow t_0} \frac{g.\bar{c}_2'(t) - g.\bar{c}_2'(t_0)}{t - t_0} = h_2.\tilde{c}''(t_0).$$

It follows that there is a $h := h_1 h_2^{-1} \in G_{\bar{c}_1'(t_0)}$ with $\bar{c}_1''(t_0) = hg.\bar{c}_2''(t_0)$, which shows the assertion.

Let E be the set of accumulation points of $c^{-1}(0)$. For connected components of $\mathbb{R}\backslash E$ we can proceed inductively to obtain twice differentiable lifts on them.

Let $\hat{c} : \mathbb{R} \to V$ be a global C^1 lift of c which exists by theorem 11.4.2 and theorem 12.3.3. We define the following set

$$F := \{t \in \mathbb{R} : \hat{c}(t) = \hat{c}'(t) = 0\}.$$

We have seen in the proof of theorem 11.4.2 that every lift $\bar{c}$ of c has to vanish and is differentiable at any point $t' \in E$ with derivative 0.

In particular this shows that $E \subseteq F$. If we denote by F' the accumulation points of F, then $E \subseteq F = (F\backslash F') \cup F' \subseteq c^{-1}(0)$.

Consider first some $t' \in F\backslash F'$, i.e., t' is an isolated point of F. Then again we have a local twice differentiable lift for $t \neq t'$ (left and right of t'), since near t' there are only points of $\mathbb{R}\backslash E$. Moreover, proposition 12.4.1 yields again a local C^1

lift near t' which is twice differentiable at t'. As above we are able to find a twice differentiable lift on the set $(\mathbb{R}\backslash E) \cup (F\backslash F')$.

Finally let $t' \in F'$, i.e., t' is an accumulation point of F. By proposition 12.4.1, we have again a local C^1 lift $\tilde{c}$ near t' which is twice differentiable at t'. Lemma 12.1.4 implies that locally near t' the set F is given by $F = \{\tilde{c}(t) = \tilde{c}'(t) = 0\}$. So we have $\tilde{c}(t') = \tilde{c}'(t') = \tilde{c}''(t') = 0$, as t' is an accumulation point of F. We extend our twice differentiable lift $\bar{c}$ on $(\mathbb{R}\backslash E) \cup (F\backslash F')$ by 0 on F' to the whole of $(\mathbb{R}\backslash E) \cup (F\backslash F') \cup F' = (\mathbb{R}\backslash E) \cup F = \mathbb{R}$. It remains to check that then $\bar{c}$ is twice differentiable at $t' \in F'$. Since $F' \subseteq E$, we obtain that $\bar{c}$ vanishes at t' and is continuous and differentiable there with derivative 0. Consider a sequence $(t_m)_m$ with $t' \neq t_m \to t'$. Passing to subsequences, we find as above, for all m, $\bar{c}(t_m) = g.\tilde{c}(t_m)$ and $\bar{c}'(t_m) = hg.\tilde{c}'(t_m)$ for some $g \in G$ and some $h \in G_{\bar{c}(t_m)}$. Then,

$$\frac{\bar{c}'(t_m) - \bar{c}'(t')}{t_m - t'} = \frac{\bar{c}'(t_m)}{t_m - t'} = \frac{hg.\tilde{c}'(t_m)}{t_m - t'} \to hg.\tilde{c}''(t') = 0 \qquad \text{as } t_m \to t'.$$

It follows that the second derivative of $\bar{c}$ at t' exists and equals 0. This completes the proof. □

12.5. Generalization to polar representations

The main results of sections 12.3 and 12.4, obtained there for finite groups G, can be generalized to polar representations $G \to \mathrm{O}(V)$.

We use the notation of section 10.3.

12.5.1. Proposition. *Let $\rho : G \to \mathrm{O}(V)$ be a polar representation of a compact Lie group on a finite dimensional Euclidean vector space V with section Σ. If Σ' is a different section, then there is an isomorphism $W(\Sigma) \to W(\Sigma')$ induced by an inner automorphism of G. It is uniquely determined up to an inner automorphism of $W(\Sigma)$.*

Proof. We have $\Sigma' = g.\Sigma$, where $g \in G$ is uniquely determined up to $N(\Sigma)$. Clearly, conjugation by g, $\mathrm{conj}_g : G \to G$, induces isomorphisms $N(\Sigma) \to N(\Sigma')$ and $Z(\Sigma) \to Z(\Sigma')$. Therefore it factors to an isomorphism $W(\Sigma) \to W(\Sigma')$. □

12.5.2. Theorem. *Let $\rho : G \to \mathrm{O}(V)$ be a polar representation of a compact Lie group on a finite dimensional Euclidean vector space V with orbit map $\sigma : V \to \mathbb{R}^n$. Assume that $W(\Sigma) \to \mathrm{O}(\Sigma)$ has property $(\mathcal{B}_k)$ for some section Σ (and hence for any section by proposition 12.5.1). Let $c : \mathbb{R} \to \sigma(V) \subseteq \mathbb{R}^n$ be a curve in the orbit space. Then we have:*

1. *If c is of class C^{k+d}, then there exists a global orthogonal C^1 lift $\bar{c} : \mathbb{R} \to V$.*
2. *If c is of class C^{k+2d}, then there exists a global orthogonal twice differentiable lift $\bar{c} : \mathbb{R} \to V$.*

Proof. By theorem 10.3.2, $\sigma|_\Sigma : \Sigma \to \mathbb{R}^n$ is the orbit map for the representation $W(\Sigma) \to \mathrm{O}(\Sigma)$, and hence the orbit spaces $V/G = \sigma(V)$ and $\Sigma/W(\Sigma) = \sigma|_\Sigma(\Sigma)$ are isomorphic.

If $c : \mathbb{R} \to \sigma(V) \cong \sigma|_\Sigma(\Sigma)$ is C^{k+d}, then by theorem 11.4.2 and theorem 12.3.3 (since $W(\Sigma)$ is finite) there exists a global C^1 lift $\bar{c} : \mathbb{R} \to \Sigma$, which as a curve in V is orthogonal to each G-orbit it meets, by the properties of Σ. This shows (1).

If $c : \mathbb{R} \to \sigma(V) \cong \sigma|_\Sigma(\Sigma)$ is C^{k+2d}, then statement (2) follows analogously from theorem 12.4.1. □

Example. The standard representation of $\mathrm{SO}(n)$ on $\mathbb{R}^n$ is polar. Any 1-dimensional linear subspace Σ of $\mathbb{R}^n$ is a section. The associated generalized Weyl group is $W(\Sigma) = \{\pm\,\mathrm{id}\}$. So the representation $W(\Sigma) \to \mathrm{O}(\Sigma)$ has property $(\mathcal{B}_2)$, since it reduces to the standard representation of S_2 on $\mathbb{R}^2$, the problem of finding regular roots of $x^2 - f(t) = 0$ ($f \geq 0$ and C^2), which has property $(\mathcal{B}_2)$, by proposition 4.1.1. Hence theorem 12.5.2 is applicable.

12.6. Polar representations have property $(\mathcal{B})$

12.6.1. THEOREM. *Let $\rho : G \to \mathrm{O}(V)$ be a real finite dimensional orthogonal representation of a finite group G, and let $\sigma_1, \ldots, \sigma_n$ be a minimal system of homogeneous generators of $\mathbb{R}[V]^G$. Write $V = V_1 \oplus \cdots \oplus V_l$ as orthogonal direct sum of irreducible subspaces V_i. Choose $v_i \in V_i \backslash \{0\}$ such that the cardinality of the corresponding isotropy group G_{v_i} is maximal, and put $k = \max\{d(\rho), |G|/|G_{v_i}| : 1 \leq i \leq l\}$. Then any curve $c = (c_1, \ldots, c_n) : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ of class C^k in the orbit space admits a global differentiable lift $\bar{c}$ to V with locally bounded derivative.*

PROOF. We shall reduce to the polynomial case, i.e., to the standard representation of S_k on $\mathbb{R}^k$.

Since $d \leq k$, we can apply theorem 11.4.2 which provides a differentiable lift $\bar{c} : \mathbb{R} \to V$ of c.

Let i be fixed. For $g \in G$ we define a linear function

$$\begin{aligned} F_{i,g} : V &\longrightarrow \mathbb{R} \\ x &\longmapsto \langle v_i \mid g.\operatorname{pr}_{V_i}(x)\rangle = \langle v_i \mid g.x\rangle. \end{aligned}$$

Here $\operatorname{pr}_{V_i} : V \to V_i$ is the natural projection. The cardinality of distinct functions $F_{i,g}$ equals $k_i := |G|/|G_{v_i}|$.

Let $G_{v_i}\backslash G$ denote the space of right cosets of G_{v_i} in G, and introduce a numbering $G_{v_i}\backslash G = \{g_1, g_2, \ldots, g_{k_i}\}$. We construct the following polynomials on V:

$$a_{i,j}(x) = \sum_{1 \leq m_1 < \cdots < m_j \leq k_i} F_{i,g_{m_1}}(x) \cdots F_{i,g_{m_j}}(x) \qquad 1 \leq j \leq k_i.$$

These polynomials $a_{i,j}$ are G-invariant by construction, and therefore expressible in the homogeneous generators $\sigma_1, \ldots, \sigma_n$ of $\mathbb{R}[V]^G$, i.e., there exist polynomials $p_{i,j} \in \mathbb{R}[\mathbb{R}^n]$ such that

$$a_{i,j} = p_{i,j}(\sigma_1, \ldots, \sigma_n) \qquad 1 \leq j \leq k_i. \tag{12.2}$$

The polynomials $a_{i,j}$, for $1 \leq j \leq k_i$, are elementary symmetric functions in the variables $F_{i,g}(x)$, where g runs through $G_{v_i}\backslash G = \{g_1, g_2, \ldots, g_{k_i}\}$. Finally, we associate the following monic polynomial of degree k_i in one variable y:

$$P_i(x)(y) = y^{k_i} + \sum_{j=1}^{k_i} (-1)^j a_{i,j}(x) y^{k_i - j} = \prod_{j=1}^{k_i} (y - F_{i,g_j}(x)). \tag{12.3}$$

By construction, the functions $x \mapsto F_{i,g}(x)$ ($g \in G_{v_i}\backslash G$) parameterize the roots of $x \mapsto P_i(x)(y)$ which, consequently, are always real.

Now consider the functions $t \mapsto a_{i,j}(\bar{c}(t))$ ($1 \leq j \leq k_i$) which are of class C^k by equation (12.2). As in (12.3) we may associate a C^k curve $t \mapsto P_i(t)(y)$ of monic polynomials of degree k_i in one variable defined by

$$P_i(t)(y) = y^{k_i} + \sum_{j=1}^{k_i} (-1)^j a_{i,j}(\bar{c}(t)) y^{k_i - j}.$$

By theorem 5.5.13 or theorem 6.3.1, applied to the curve of polynomials $t \mapsto P_i(t)(y)$, the differentiable functions $t \mapsto F_{i,g}(\bar{c}(t))$ ($g \in G_{v_i}\backslash G$) which parameterize the roots of $t \mapsto P_i(t)(y)$ have locally bounded derivative.

Since V_i is irreducible, the linear span of the orbit $G.v_i$ spans V_i. If we repeat the above procedure for each $1 \leq i \leq l$, it follows that $\bar{c}$ is a differentiable lift of c with locally bounded derivative. This completes the proof. □

COROLLARY. *Any real finite dimensional orthogonal representation $\rho : G \to \mathrm{O}(V)$ of a finite group G has property $(\mathcal{B}_k)$, where $k = \max\{d(\rho), |G|/|G_{v_i}| : 1 \leq i \leq l\}$, $v_i \in V_i \backslash \{0\}$ are chosen such that the cardinality of G_{v_i} is maximal, and $V = V_1 \oplus \cdots \oplus V_l$ is the decomposition into irreducible subrepresentations.* □

12.6.2. COROLLARY. *Any polar representation ρ of a compact Lie group G has property $(\mathcal{B}_k)$, where k is determined analogously to corollary 12.6.1 but for the representation $W \to \mathrm{O}(\Sigma)$, where W is the generalized Weyl group of some section Σ. Moreover, the lifts can be chosen orthogonal.* □

12.6.3. REMARKS. (1) The case $k = |G|$ can occur: For finite rotation groups in the plane we have $d = |G|$, and for any non-zero v the isotropy group G_v is trivial.

(2) There are irreducible orthogonal representations of finite groups G where the inequality $d \leq |G|/|G_v|$ is violated for non-zero vectors v: Consider the rotational symmetry group T of the regular tetrahedron in $\mathbb{R}^3$. We find $d = 6$ (e.g. [**CP88**]). The isotropy group of each vertex v of the tetrahedron has 3 elements. So $|G|/|G_v| = 12/3 = 4$.

Furthermore, the same phenomenon appears for the rotational symmetry groups W and H of the cube and the regular icosahedron in $\mathbb{R}^3$, respectively. See section 12.8.

12.7. Property $(\mathcal{B}_k)$ for finite reflection groups

12.7.1. Finite reflection groups. Let V be a finite dimensional Euclidean vector space. A finite subgroup G of $\mathrm{O}(V)$ generated by *reflections*, i.e., linear transformations $V \to V$ that carry each vector to its mirror image with respect to a fixed hyperplane, is called *finite reflection group*. A finite effective reflection group is called *Coxeter group*. Suppose that $S \in G$ is a reflection through a hyperplane P, then the two unit vectors that are perpendicular to P are called *roots* of G. The set Δ of all roots corresponding to the generating reflections, together with all images of this roots under all transformations of G is called a *root system* for G. Choose a vector $t \in V$ such that $\langle t \mid r\rangle \neq 0$ for each root $r \in \Delta$. Then the root system Δ is partitioned into two subsets, $\Delta_t^+ = \{r \in \Delta : \langle t \mid r\rangle > 0\}$ and $\Delta_t^- = \{r \in \Delta : \langle t \mid r\rangle < 0\}$. Choose a subset Π of Δ_t^+ that is minimal with respect to the property that every $r \in \Delta_t^+$ is a linear combination, with all coefficients non-negative, of elements of Π. We say that Π is a t*-base* of Δ. One can show that it is unique. The roots in the t-base Π are called *simple roots* , the reflections along simple roots *simple reflections*. The simple reflections generate G, and, if for $T \in G$ we have $T(\Pi) = \Pi$, then $T = \mathrm{id}$, e.g. [**GB85**].

12.7.2. Fundamental domains. An open subset F of V is called *fundamental domain* if and only if $F \cap T(F) = \emptyset$ for $\mathrm{id} \neq T \in G$ and $V = \bigcup_{T\in G} \overline{T(F)}$. *We claim that the set*

$$F = F_t := \{x \in V : \langle x \mid r\rangle > 0 \textit{ for all } r \in \Pi\}$$

is a fundamental domain for the Coxeter group G. For: F is clearly open. Suppose that $T \in G$ and $x \in F \cap T(F)$. Then $T^{-1}x \in F$. Since $x \in F$, we have $\langle x \mid r\rangle > 0$ for all $r \in \Pi$, and thus $\langle x \mid s\rangle > 0$ for all $s \in \Delta_t^+$. It follows immediately that $\Delta_x^+ = \Delta_t^+$, and that $\Pi_x = \Pi_t$, by uniqueness. Analogously, we get $\Pi_{T^{-1}x} = \Pi_t$. For any $S \in G$ we have

$$S(\Delta_t^+) = \{Sr \in \Delta : \langle t \mid r\rangle = \langle St \mid Sr\rangle > 0\} = \{s \in \Delta : \langle St \mid s\rangle > 0\} = \Delta_{St}^+,$$

and hence $S(\Pi_t) = \Pi_{St}$, by uniqueness. Therefore, $\Pi_t = \Pi_{T^{-1}x} = T^{-1}(\Pi_x) = T^{-1}(\Pi_t)$. But that implies $T = T^{-1} = \mathrm{id}$. Finally, for any $y \in V$ there is a $T \in G$ such that $\langle Ty \mid r\rangle \geq 0$ for all $r \in \Pi$, and so $Ty \in \overline{F}$: Set $x_0 = \frac{1}{2}\sum_{s\in\Delta_t^+} s$ and choose $T \in G$ for which $\langle Ty \mid x_0\rangle$ is maximal; if S is the reflection along $r \in \Pi$, then one easily sees that $Sx_0 = x_0 - r$, and, hence, $\langle Ty \mid x_0\rangle \geq \langle STy \mid x_0\rangle = \langle Ty \mid Sx_0\rangle = \langle Ty \mid x_0 - r\rangle = \langle Ty \mid x_0\rangle - \langle Ty \mid r\rangle$. Consequently, $y \in T^{-1}(\overline{F}) = \overline{T^{-1}(F)}$, and the claim is proved.

12.7.3. Characterization of finite reflection groups. Let $\Pi = \{r_1, \ldots, r_n\}$ be the set of simple roots. We associate a graph Γ as follows: Let Γ have n nodes, and if $i \neq j$ the i-th and j-th nodes are joined by a branch if and only if $\langle r_i \mid r_j\rangle \neq 0$. In that case we may write $\frac{\langle r_i|r_j\rangle}{\|r_i\|\|r_j\|} = -\cos(\frac{\pi}{p_{ij}})$ for exactly one real number $p_{ij} > 2$, and the branch is labeled p_{ij}. In fact, p_{ij} is the order of S_iS_j (where S_i, S_j denote the reflections along r_i, r_j) as group element and therefore an integer, e.g. [**GB85**]. Then Γ is called *Coxeter graph* of G.

Abusing notation we will denote finite reflection groups as well as their root systems (respectively their Coxeter graphs) with the same symbols.

Recall the characterization of finite reflection groups:

Theorem ([**GB85**], [**Hum90**]). *If G is a finite subgroup of $\mathrm{O}(V)$ that is generated by reflections, then V may be written as the orthogonal direct sum of G-invariant subspaces $V_0 = V^G, V_1, \ldots, V_k$ with the following properties:*

(*a*) *If $G_i = \{g|_{V_i} : g \in G\}$, then G_i is a subgroup of $\mathrm{O}(V_i)$, and G is isomorphic with $G_0 \times G_1 \times \cdots \times G_k$.*

(*b*) *G_0 consists only of the identity transformation on V_0.*

(*c*) *Each G_i $(i \geq 1)$ is one of the groups*

$$\mathrm{A}_n, n \geq 1; \mathrm{B}_n, n \geq 2; \mathrm{D}_n, n \geq 4; \mathrm{I}_2^n, n \geq 5, n \neq 6; \mathrm{G}_2; \mathrm{H}_3; \mathrm{H}_4; \mathrm{F}_4; \mathrm{E}_6; \mathrm{E}_7; \mathrm{E}_8\,.$$

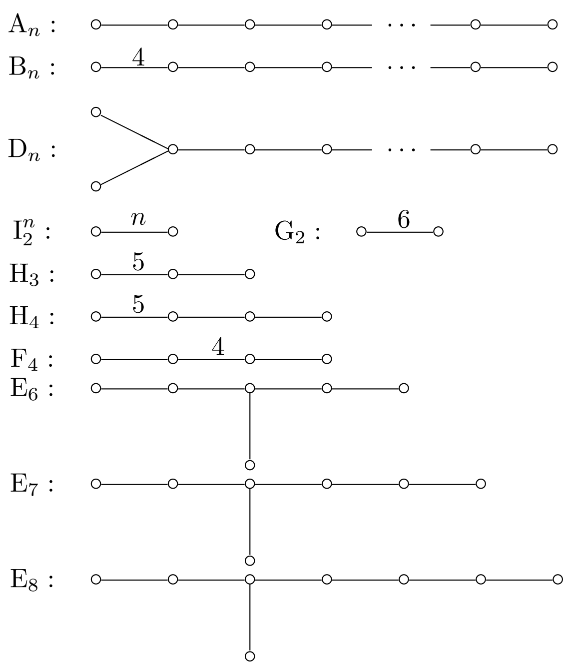

Figure 12.1. Classification of Coxeter graphs.

12.7.4. Let $\rho : G \to \mathrm{O}(V)$ be the standard representation of some irreducible finite reflection group G listed in theorem 12.7.3(c). We consider an arbitrary slice representation $G_v \to \mathrm{O}(N_v)$ ($v \in V$) of ρ; note $N_v = V$. There exists a $g \in G$ such that $g.v = w$ for a w in $\overline{F} = \{x \in V : \langle x \mid r\rangle \geq 0 \text{ for all } r \in \Pi\}$, where Π is a system of simple roots. *We assert that G_w is generated by the simple reflections it contains:* Let $T \in G_w$. By $n(T) := \mathrm{card}(\Delta^+ \cap T^{-1}(\Delta^-))$ we denote the number of positive roots sent to negative roots by T. We make induction on $n(T)$. If $n(T) = 0$, then $T = \mathrm{id}$ and we are done. If $n(T) > 0$, then T must send some simple root r to a negative root (otherwise $T(\Pi)$ and hence $T(\Delta^+)$ would consist of positive roots). Let S_r denote the reflection along r, then $n(TS_r) = n(T) - 1$ which is seen as follows: Let $r \neq s \in \Delta^+$ and write $s = \sum_{r_i \in \Pi} c_i r_i$ with $c_i \geq 0$. We have $c_k > 0$ for some $r_k \neq r$. Then $S_r s = s - 2\langle r \mid s\rangle r$ is a linear combination of Π involving r_k with the same coefficient c_k. Since $S_r s \in \Delta$ and so all coefficients must have the same sign, $S_r s \in \Delta^+$ and cannot be r (otherwise $s = S_r S_r s = S_r r = -r \in \Delta^-$). It follows that S_r maps $\Delta^+\backslash\{r\}$ into $\Delta^+\backslash\{r\}$ injectively, and hence onto $\Delta^+\backslash\{r\}$. This implies

$$\begin{aligned} S_r(\Delta^+ \cap (TS_r)^{-1}(\Delta^-)) &= S_r(\Delta^+\backslash\{r\} \cup \{r\}) \cap T^{-1}(\Delta^-) \\ &= (\Delta^+\backslash\{r\} \cup \{-r\}) \cap T^{-1}(\Delta^-) \\ &= \Delta^+\backslash\{r\} \cap T^{-1}(\Delta^-) \\ &= (\Delta^+ \cap T^{-1}(\Delta^-))\backslash\{r\}, \end{aligned}$$

since $Tr \in \Delta^-$. Thus we obtain $n(TS_r) = n(T) - 1$. Further, since $w \in \overline{F}$ and $Tr \in \Delta^-$, we find $0 \geq \langle w \mid Tr\rangle = \langle w \mid r\rangle \geq 0$, whence $\langle w \mid r\rangle = 0$ and so $S_r w = w$. Then $TS_r \in G_w$ and TS_r is a product of simple reflections contained in G_w, by induction hypothesis. Consequently, T is also such a product, and the assertion is proved.

12.7.5. Property ($\mathcal{B}_k$) for irreducible Coxeter groups. We will apply theorem 12.6.1 to each of the irreducible finite reflection groups listed in theorem 12.7.3(c). Here the inequality $d \leq |G|/|G_v|$ will be satisfied for all non-zero v which can be checked directly in figure 12.2.

It follows from 12.7.4 that we can read off easily the information we need to determine a minimal $|G|/|G_v|$ from the Coxeter graph of G. Therefore, we give a complete list of all relevant Coxeter graphs in figure 12.1.

Easy computations yield the results collected in the following figure 12.2 which gives a complete survey of the standard representations of all irreducible finite reflection groups. The integers d and $|G|$ for the listed representations can be found e.g. in [**GB85**]; in [**Meh88**] also generators of the corresponding algebra of invariant polynomials are available. See also 13.8 in the appendix. The integer k is the minimum of the numbers $|G|/|G_v|$ where v runs through all non-zero vectors in V. By theorem 12.6.1 the representations listed in the table have property ($\mathcal{B}_k$). This together with lemma 10.1.2 and proposition 12.2.1 treats finite reflection groups completely.

12.8. Property ($\mathcal{B}_k$) for finite rotation groups

12.8.1. C_2^n has property ($\mathcal{B}_n$). Let us denote by C_2^n the cyclic subgroup of $\mathrm{O}(2)$ generated by the counterclockwise rotation of $\mathbb{R}^2$ through the angle $2\pi/n$. Here we have $d = |G| = n$, and for any non-zero vector $v \in \mathbb{R}^2$ its isotropy group G_v

$\rho : G \to \mathrm{O}(V)$	d	k	$\lvert G\rvert$
A_n, $n \geq 1$	$n+1$	$n+1$	$(n+1)!$
B_n, $n \geq 2$	$2n$	$2n$	$2^n n!$
D_n, $n \geq 4$	$2n-2$	$2n$	$2^{n-1}n!$
I_2^n, $n \geq 5$	n	n	$2n$
G_2	6	6	12
H_3	10	12	120
H_4	30	120	14400
F_4	12	24	1152
E_6	12	27	51840
E_7	18	56	2903040
E_8	30	240	696729600

FIGURE 12.2. Irreducible Coxeter groups and associated integers d, k, and $|G|$.

is trivial. By corollary 12.6.1, finite rotation groups C_2^n in the plane have property $(\mathcal{B}_n)$.

REMARK. This together with the result for dihedral groups I_2^n in section 12.7 gives a complete discussion of all finite subgroups of $\mathrm{O}(2)$.

12.8.2. Finite rotation groups in three dimensions. Next we consider finite rotation groups in 3-dimensional space.

Let P be a 2-dimensional linear subspace in $\mathbb{R}^3$. Any rotation R in $\mathrm{O}(P)$ can be extended to a rotation in $\mathrm{O}(3)$, by setting $Rx = x$ for all $x \in P^\perp$ and using linearity. By extending each transformation in a cyclic subgroup C_2^n of $\mathrm{O}(P)$ in this fashion, we obtain a cyclic subgroup of rotations in $\mathrm{O}(3)$, which will be denoted by C_3^n.

On the other hand, if S is a reflection in $\mathrm{O}(P)$, then S may also be extended to a rotation in $\mathrm{O}(3)$, in fact to the rotation through the angle π having the reflection line of S in P as its axis of rotation: Define $Sx = -x$ for all $x \in P^\perp$ and extend by linearity. By extending each transformation in a dihedral subgroup I_2^n of $\mathrm{O}(P)$ to a rotation in $\mathrm{O}(3)$, the resulting set of rotations is a subgroup of $\mathrm{O}(3)$ isomorphic with I_2^n; it shall be denoted I_3^n.

Let T, W, and H denote the subgroups of rotations in $\mathrm{O}(3)$ which leave invariant the regular tetrahedron, cube, and icosahedron each with center in the origin.

THEOREM (e.g. [**GB85**]). *The following list provides a complete characterization of finite rotation groups in* $\mathbb{R}^3$*:*

$$\mathrm{C}_3^n, n \geq 1; \mathrm{I}_3^n, n \geq 2; \mathrm{T}; \mathrm{W}; \mathrm{H}\,.$$

12.8.3. C_3^n has property $(\mathcal{B}_n)$. This is by construction, since the linear subspace $P^\perp$ is left pointwise invariant under the C_3^n-action, and on P it restricts to the C_2^n-action; so lemma 10.1.2 and the result for C_2^n give the statement.

12.8.4. I_3^n has property $(\mathcal{B}_{n+1})$. Note first that here $d = n+1$. Moreover, we have the decomposition $\mathbb{R}^3 = P \oplus P^\perp$ into irreducible subrepresentations. In order to make $|G|/|G_{v_1}|$ minimal for $0 \neq v_1 \in P$ we may choose v_1 to lie on some reflection line on I_2^n in P; then $|G|/|G_{v_1}| = 2n/2 = n$. For $0 \neq v_2 \in P^\perp$ we find that $|G_{v_2}|$ is the number of rotations (including the identity) in I_2^n. So $|G|/|G_{v_2}| = 2n/n = 2$. Application of corollary 12.6.1 gives the assertion.

12.8.5. T **has property** ($\mathcal{B}_6$)**.** We have $d = 6$ (see 13.8, or [**CP88**], [**PSW78**]). Further the action of T on $\mathbb{R}^3$ is irreducible. The elements of T consist of rotations through angles of $2\pi/3$ and $4\pi/3$ about each of four axes joining vertices of the tetrahedron with centers of opposite faces, rotations through the angle π about each of the three axes joining the midpoints of opposite edges, and the identity. So $|\,\mathrm{T}\,| = 12$. The isotropy groups of non-zero vectors on axes joining vertices with centers of opposite faces have cardinality 3, those of non-zero vectors on axes joining the midpoints of opposite edges have cardinality 2, and all other isotropy groups of non-zero vectors are trivial. Hence application of corollary 12.6.1 gives the statement.

12.8.6. W **has property** ($\mathcal{B}_9$)**.** We have $d = 9$ (see 13.8 or [**PSW78**]). The action of W on $\mathbb{R}^3$ is irreducible. The elements of W consist of rotations through angles of $\pi/2$, π, and $3\pi/2$ about each of three axes joining the centers of opposite faces, rotations through angles of $2\pi/3$ and $4\pi/3$ about each of four axes joining extreme opposite vertices, rotations through the angle π about each of six axes joining midpoints of diagonally opposite edges, and the identity. Thus $|\,\mathrm{W}\,| = 24$. The isotropy groups of non-zero vectors on axes joining the centers of opposite faces have cardinality 4, those of non-zero vectors on axes joining extreme opposite vertices have cardinality 3, those of non-zero vectors on axes joining midpoints of diagonally opposite edges have cardinality 2, and all other isotropy groups of non-zero vectors are trivial. Apply corollary 12.6.1.

12.8.7. H **has property** ($\mathcal{B}_{15}$)**.** We have $d = 15$ (see 13.8 or [**CP88**]). The action of H on $\mathbb{R}^3$ is irreducible. The elements of H consist of rotations through angles of $2\pi/5$, $4\pi/5$, $6\pi/5$, and $8\pi/5$ about each of the six axes joining extreme opposite vertices, rotations through angles of $2\pi/3$ and $4\pi/3$ about each of ten axes joining centers of opposite faces, rotations through the angle π about each of fifteen axes joining midpoints of opposite edges, and the identity. Therefore $|\,\mathrm{H}\,| = 60$. The isotropy groups of non-zero vectors on axes joining extreme opposite vertices have cardinality 5, those of non-zero vectors on axes joining centers of opposite faces have cardinality 3, those of non-zero vectors on axes joining midpoints of opposite edges have cardinality 2, and all other isotropy groups of non-zero vectors are trivial. Apply corollary 12.6.1.

12.8.8. The following figure 12.3 collects the results for finite rotation groups in two and three dimensions obtained in this section. The groups in the first column of the table are meant to stay for their standard representation, d is the integer associated to representations in section 11.2, and k is as in corollary 12.6.1.

$\rho : G \to \mathrm{O}(V)$	d	k	$\|G\|$
C_2^n, $n \geq 1$	n	n	n
C_3^n, $n \geq 1$	n	n	n
I_3^n, $n \geq 2$	$n+1$	$n+1$	$2n$
T	6	6	12
W	9	9	24
H	15	15	60

FIGURE 12.3. Finite rotation groups in two and three dimensions and associated integers d, k, and $|G|$.

REMARK. Observe that in this table we always have $d = k$, i.e., the respective representation has property $(\mathcal{B}_d)$. Since we need at least regularity C^d for a curve in the orbit space to be liftable once differentiably (theorem 11.4.2), we cannot expect to improve these results. Evidently this remark applies also for those representations in the table in section 12.7 with $d = k$.

12.9. Lipschitz lifts of mappings

12.9.1. LEMMA. *Let $\rho : G \to \mathrm{O}(V)$ be a real finite dimensional representation of a finite group G. Consider a curve $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ of class C^d, where d is defined as in theorem 12.6.1. Then any continuous lift $\bar{c} : \mathbb{R} \to V$ possesses one-sided first derivatives at any $t \in \mathbb{R}$.*

PROOF. By theorem 11.4.2, there exists a global differentiable lift $\tilde{c}$ of c. Let $(t_m)_m \subseteq \mathbb{R}$ a sequence with $t_m \searrow 0$. Then, we find $g_m \in G$ such that $\bar{c}(t_m) = g_m.\tilde{c}(t_m)$, for all m. By passing to a subsequence (again denoted by $(t_m)_m$), we obtain $\bar{c}(t_m) = g.\tilde{c}(t_m)$, for all m and fixed $g \in G$. We find

$$\lim_{t_m \searrow 0} \frac{\bar{c}(t_m)}{t_m} = \lim_{t_m \searrow 0} \frac{g.\tilde{c}(t_m)}{t_m} = g.\tilde{c}'(0).$$

It follows that the accumulation points of $\frac{\bar{c}(t)}{t}$ as $t \searrow 0$ lie in the orbit $G.\tilde{c}'(0)$. By lemma 11.4.1, the right-sided derivative $\bar{c}^{(+)}(0) = \lim_{t \searrow 0} \frac{\bar{c}(t)}{t}$ exists. Similarly, for the left-sided derivative. ☐

12.9.2. THEOREM. *Let V be a real finite dimensional vector space and let $G \subseteq \mathrm{O}(V)$ be a finite subgroup generated by reflections. Let k be defined as in theorem 12.6.1. Any C^k mapping $f : \mathbb{R}^m \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ allows a continuous lift $\bar{f} : \mathbb{R}^m \to V$ which is locally Lipschitz. Actually, any continuous lift $\bar{f}$ is locally Lipschitz.*

PROOF. The mapping f allows a continuous lift $\bar{f}$ to V, since the orbit space $V/G = \sigma(V)$ is homeomorphic to the closure of a fundamental domain sitting in V.

We shall show that $f \circ l : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ has the property that any continuous lift $\widetilde{f \circ l} : \mathbb{R} \to V$ is locally Lipschitz, for any line l parallel to one of the coordinate axes and parameterized with constant speed. This will imply the statement of the theorem.

Let us assume that $f : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ is of class C^k. By theorem 12.6.1, there exists a global differentiable lift $\tilde{f}$ of f with locally bounded derivative. By lemma 12.9.1, for any continuous lift $\bar{f}$ the one-sided derivatives $\bar{f}^{\pm}(t)$ exist for all t, and, by lemma 12.1.4, the differ from $\tilde{f}'(t)$ only by permutations. Hence any continuous lift $\bar{f}$ of t is locally Lipschitz. ☐

12.9.3. COROLLARY. *Let $\rho : G \to \mathrm{O}(V)$ be a polar representation of a compact connected Lie group G. Let $\sigma_1, \ldots, \sigma_n$ a minimal system of homogeneous generators of the algebra of invariant polynomials $\mathbb{R}[V]^G$. Let k be as in corollary 12.6.2. Any C^k mapping $f : \mathbb{R}^m \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ allows a continuous lift $\bar{f} : \mathbb{R}^m \to V$ which is locally Lipschitz.*

PROOF. Let Σ be a section and let $W = W(\Sigma)$ be the associated generalized Weyl group. Since G is connected, W is a finite reflection group ([**DK85**]). The restriction $\sigma|_\Sigma : \Sigma \to \mathbb{R}^n$ is an orbit map for the W-action on Σ, and the orbit spaces $V/G = \sigma(V)$ and $\Sigma/W = \sigma|_\Sigma(\Sigma)$ are isomorphic.

By theorem 12.9.2, a C^k mapping $f : \mathbb{R}^m \to \sigma(V) \cong \sigma|_\Sigma(\Sigma)$ admits a continuous lift $\bar{f} : \mathbb{R}^m \to \Sigma$ which is locally Lipschitz. ☐

CHAPTER 13

Choosing roots of polynomials with symmetries smoothly

Consider a smooth curve of monic hyperbolic polynomials with fixed degree n:

$$P(t)(x) = x^n - a_1(t)x^{n-1} + a_2(t)x^{n-2} - \cdots + (-1)^n a_n(t) \qquad (t \in \mathbb{R}).$$

We have seen in theorem 4.3.1 that the roots of $P(t)$ may be chosen smoothly if no two of the increasingly ordered continuous roots meet of infinite order. In general, as shown in theorem 7.1.1, the roots of a C^{3n} curve $P(t)$ of hyperbolic polynomials can be parameterized twice differentiable (and not better).

The space Hyp_n of monic hyperbolic polynomials P of fixed degree n may be identified with a semialgebraic subset in $\mathbb{R}^n$, the coefficients of P being the coordinates. Then $P(t)$ is a smooth curve in $\mathrm{Hyp}_n \subseteq \mathbb{R}^n$. If the curve $P(t)$ lies in some semialgebraic subset of Hyp_n, then it is evident that in general the conditions which guarantee smooth parameterizations of the roots of $P(t)$ are weaker than those mentioned in the previous paragraph. We are going to study that phenomenon.

It will turn out that under the assumption that the polynomials $P(t)$ have certain symmetries we can indeed improve the conditions for regular parameterizations of the roots. These improvements are essentially applications of the lifting problem tackled in chapters 10, 11, and 12.

We follow [**LR07**].

13.1. Preliminaries

In this chapter we shall denote by E_j the elementary symmetric functions

$$E_j(x) = \sum_{1 \le i_1 < \cdots < i_j \le n} x_{i_1} \cdots x_{i_j} \qquad (1 \le j \le n).$$

The first n Newton polynomials

$$N_i(x_1, \ldots, x_n) = \sum_{j=1}^{n} x_j^i$$

constitute a different system of generators of $\mathbb{R}[\mathbb{R}^n]^{\mathrm{S}_n}$. In this notation (3.1) amounts to

$$N_k - N_{k-1}E_1 + N_{k-2}E_2 + \cdots + (-1)^{k-1}N_1E_{k-1} + (-1)^k kE_k = 0 \quad (k \ge 1). \tag{13.1}$$

For convenience we shall switch from elementary symmetric functions to Newton polynomials and conversely, if it seems appropriate.

13.1.1. Lemma. *The equivalent conditions* (1) *and* (2) *in theorem 4.3.1 are satisfied at t_0 if and only if P is normally nonflat at t_0 as curve in $E(\mathbb{R}^n) = \mathbb{R}^n/\,\mathrm{S}_n$.*

Proof. Let P be normally nonflat at t_0. We use the notation of 10.1.6. Let s be a minimal integer such that $P(t)$ lies in A_s for t near t_0 and let $f \in I_s$ be such that $f \circ P$ is not infinitely flat at t_0. Denote by $\bar{I}_s$ the ideal in $\mathbb{R}[\mathbb{R}^n]$ defining the closed subset $\pi^{-1}(A_{s-1}) \subseteq \mathbb{R}^n$, where $\pi : \mathbb{R}^n \to \mathbb{R}^n/\,\mathrm{S}_n$ is the quotient projection. It is easy to see that the polynomials

$$f_{i_1\dots i_s} = (x_{i_1} - x_{i_2}) \cdots (x_{i_1} - x_{i_s}) \cdots (x_{i_{s-1}} - x_{i_s}),$$

where $1 \le i_1 < \cdots < i_s \le n$, generate $\bar{I}_s$. So there exist polynomials $Q_{i_1\dots i_s} \in \mathbb{R}[\mathbb{R}^n]$ such that

$$f \circ \pi = \sum_{i_1<\cdots<i_s} Q_{i_1\dots i_s} f_{i_1\dots i_s}.$$

Denote by $\bar{P}(t)$ the lift of $P(t)$ given by the increasingly ordered continuous roots $x_1(t), \dots, x_n(t)$ of the polynomial $P(t)$. Then we have

$$f \circ P(t) = \sum_{i_1<\cdots<i_s} Q_{i_1\dots i_s} \circ \bar{P}(t) \cdot f_{i_1\dots i_s} \circ \bar{P}(t).$$

Since $f \circ P$ is not infinitely flat at t_0, at least one of the summands in this sum is not infinitely flat at t_0 and thus there is a polynomial $f_{i_1\dots i_s}$ such that $f_{i_1\dots i_s} \circ \bar{P}$ is not infinitely flat at t_0. By assumption, among the roots $x_1(t), \dots, x_n(t)$ there are precisely s distinct for t near t_0. Hence the germs at t_0 of the roots $x_{i_1}(t), \dots, x_{i_s}(t)$ are distinct, and no two of them meet of infinite order at t_0. Therefore, condition (1) in theorem 4.3.1 is satisfied.

The other direction is evident by (3.3). □

13.2. Lifting smooth curves in spaces of hyperbolic polynomials

13.2.1. The problem. Let us denote by Hyp_n the space of hyperbolic polynomials of degree n

$$P(x) = x^n + \sum_{j=1}^{n} (-1)^j a_j x^{n-j}.$$

We may naturally view Hyp_n as a semialgebraic subset of $\mathbb{R}^n$ by identifying P with $(a_1, \dots, a_n)$. We have $\mathrm{Hyp}_n = E(\mathbb{R}^n) = \mathbb{R}^n/\,\mathrm{S}_n$, and, by means of theorem 3.1.2, we may calculate explicitly a set of inequalities defining Hyp_n.

Suppose X is a semialgebraic subset of Hyp_n. Let $c : \mathbb{R} \to X$ be a smooth curve in X; smooth as curve in $\mathbb{R}^n$. We may view c as a curve in Hyp_n, i.e., as a smooth curve of monic hyperbolic polynomials of degree n. In theorem 4.3.1 sufficient conditions for the existence of a smooth lift $\bar{c}$ to $\mathbb{R}^n$, i.e., a smooth parameterization of its roots, are presented. It is evident that a smooth curve c in X in order to be liftable smoothly over E to $E^{-1}(X)$ must in general fulfill weaker genericity conditions. Our purpose is to investigate that phenomenon.

13.2.2. Orbit spaces embedded in spaces of hyperbolic polynomials. We recall a construction due to L. Smith and R.E. Strong [**SS87**] (see also [**BR83**]) related to E. Noether's [**Noe15**] proof of Hilbert's finiteness theorem as recounted by H. Weyl [**Wey39**].

Let $\rho : G \to \mathrm{GL}(V)$ be a representation of a finite group G in a finite dimensional vector space V. Consider its induced representation in the dual V^*. For an orbit $B \subseteq V^*$ set

$$\phi_B(X) = \prod_{b\in B} (X + b)$$

which we regard as an element of the ring $\mathbb{R}[V][X]$, with X a new variable. The polynomial $\phi_B(X)$ is called the *orbit polynomial* of B. Evidently, $\phi_B \in \mathbb{R}[V]^G[X]$. If $|B|$ denotes the cardinality of the orbit B, we may expand $\phi_B(X)$ to a polynomial of degree $|B|$ in X,

$$\phi_B(X) = \sum_{i+j=|B|} C_i(B) X^j,$$

defining classes $C_i(B) \in \mathbb{R}[V]^G$ called the *orbit Chern classes* of B.

THEOREM (L. Smith and R.E. Stong [**SS87**]). *Let $\rho : G \hookrightarrow \mathrm{GL}(V)$ be a faithful representation of a finite group G. Then there exist orbits $B_1, \dots, B_l \subseteq V^*$ such that the associated orbit Chern classes $C_i(B_j)$, $1 \le i \le |B_j|$, $1 \le j \le l$, generate $\mathbb{R}[V]^G$.*

The field of real numbers may be replaced by any field of either characteristic zero or characteristic larger than the order of G. For our purpose the reals will suffice.

The Chern classes of the orbit are exactly the elementary symmetric functions in the elements of the orbit. If $B \subseteq V^*$ is an orbit and V_B^* is a vector space with basis identified with the elements of B, then there is a natural map $V_B^* \to V^*$ given by the identification. This map induces a map $\mathbb{R}[V_B]^{\mathrm{S}_{|B|}} \to \mathbb{R}[V]^G$ which sends the k-th elementary symmetric function to the k-th orbit Chern class of B.

In this notation the above theorem says that there exist orbits $B_1, \dots, B_l \subseteq V^*$ such that the induced map

$$\bigotimes_{i=1}^{l} \mathbb{R}[V_{B_i}]^{\mathrm{S}_{|B_i|}} \longrightarrow \mathbb{R}[V]^G$$

is surjective.

The orbit Chern classes $C_i(B)$ of an orbit B, viewed as invariant polynomials on V, define a G-invariant map

$$C(B) = (C_1(B), \dots, C_{|B|}(B)) : V \longrightarrow \mathbb{R}^{|B|}$$

whose image $C(B)(V)$ is a semialgebraic subset of the space $\mathrm{Hyp}_{|B|}$ of hyperbolic polynomials of degree $|B|$.

By the above theorem, for any faithful representation $\rho : G \hookrightarrow \mathrm{GL}(V)$ of a finite group G there exist orbits $B_1, \dots, B_l \subseteq V^*$ such that the map

$$C(\rho) = (C(B_1), \dots, C(B_l)) : V \longrightarrow \mathrm{Hyp}_{|B_1|} \times \dots \times \mathrm{Hyp}_{|B_l|} \subseteq \mathbb{R}^{|B_1| + \dots + |B_l|}$$

induces a homeomorphism between the orbit space V/G and the image $C(\rho)(V)$ which is a semialgebraic subset of $\mathrm{Hyp}_{|B_1|} \times \dots \times \mathrm{Hyp}_{|B_l|}$. By increasing the number of orbits B_i if necessary, we may assume that each irreducible subspace of V contributes at least one orbit B_i. Then, the linear forms $b \in B_1 \cup \dots \cup B_l$ induce an injective inclusion $V \hookrightarrow \mathbb{R}^{|B_1| + \dots + |B_l|}$.

Let $c : \mathbb{R} \to C(\rho)(V)$ be a smooth curve. Then $c = (c_1, \dots, c_l)$ where each $c_i : \mathbb{R} \to C(B_i)(V)$ is smooth. Since $C(B_i)(V) \subseteq \mathrm{Hyp}_{|B_i|}$ we may view c_i as a curve in $\mathrm{Hyp}_{|B_i|}$. If there exist smooth lifts $\bar{c}_i : \mathbb{R} \to \mathbb{R}^{|B_i|}$ with respect to the representations $\mathrm{S}_{|B_i|} : \mathbb{R}^{|B_i|}$, then $\bar{c} = (\bar{c}_1, \dots, \bar{c}_l) : \mathbb{R} \to \mathbb{R}^{|B_1| + \dots + |B_l|}$ is a smooth lift with respect to $\mathrm{S}_{|B_1|} \times \dots \times \mathrm{S}_{|B_l|} : \mathbb{R}^{|B_1| + \dots + |B_l|}$. Consequently, it suffices to study the case when there is given a smooth curve in a semialgebraic subset of some Hyp_n. That is exactly the problem introduced in 13.2.1.

Suppose $\tilde{c} : \mathbb{R} \to V$ is a smooth lift of c with respect to ρ. Then, there exists a smooth lift $\bar{c} : \mathbb{R} \to \mathbb{R}^{|B_1| + \dots + |B_l|}$ of c with respect to the representation of $\mathrm{S}_{|B_1|} \times \dots \times \mathrm{S}_{|B_l|}$ on $\mathbb{R}^{|B_1| + \dots + |B_l|}$, namely

$$\begin{tikzcd} & V \arrow[r, hook] \arrow[d] & \mathbb{R}^{|B_1| + \dots + |B_l|} \arrow[d] \\ \mathbb{R} \arrow[ur, "\tilde{c}"] \arrow[r, "c"'] & C(\rho)(V) \arrow[r, hook] & \mathrm{Hyp}_{|B_1|} \times \dots \times \mathrm{Hyp}_{|B_l|} \end{tikzcd}$$

It follows, by theorem 10.2.1, that conditions which guarantee that c is generic as curve in the orbit space V/G suffice to imply the existence of a smooth lift of c with respect to $\mathrm{S}_{|B_1|} \times \cdots \times \mathrm{S}_{|B_l|} : \mathbb{R}^{|B_1|+\cdots+|B_l|}$.

We have seen that the above construction provides a class of semialgebraic subsets of spaces of hyperbolic polynomials, namely orbit spaces of faithful finite group representations, for which we are able to apply the strategy described in 13.2.1, thanks to the result in theorem 10.2.1.

In the remaining sections we shall change the point of view. Assume we are given a curve of hyperbolic polynomials with certain symmetries. We will investigate whether we can weaken the conditions in theorem 4.3.1 which guarantee the existence of smooth parameterizations of the roots. This will be performed in section 13.4. The following section provides the necessary preparation.

13.3. Orbit type and ambient stratification

13.3.1. Orbit type and ambient stratification of U. Suppose U is a linear subspace of $\mathbb{R}^n$. Let the symmetric group S_n act on $\mathbb{R}^n$ by permuting the coordinates and endow U with the induced effective action of

$$W = W(U) := N(U)/Z(U),$$

where $N(U) := \{\tau \in \mathrm{S}_n : \tau.U = U\}$ and $Z(U) := \{\tau \in \mathrm{S}_n : \tau.x = x \text{ for all } x \in U\}$. Then U carries two natural stratifications, namely, the orbit type stratification with respect to the W-action and the restriction to U of the orbit type stratification of $\mathbb{R}^n$ with respect to the S_n-action. It is easily seen that the latter indeed provides a Whitney stratification of U. Let us denote it as the *ambient stratification of U*.

Remark. Note that in view of definition 9.6.4 we should rather speak of ambient *decomposition* than stratification. But we maintain that abuse of notation, since there will be no consequences for our purpose.

13.3.2. Proposition. *Let U be a linear subspace in $\mathbb{R}^n$ endowed with the induced action by $W = W(U)$. Then for the ambient and orbit type stratification of U we have:*

(1) *Each ambient stratum is contained in a unique orbit type stratum.*
(2) *Each orbit type stratum contains at least one ambient stratum of the same dimension and is the union of all contained ambient strata.*

Proof. To (1): Let S be an ambient stratum, i.e., S is a component of $\mathrm{S}_n.\mathbb{R}^n_H \cap U$, where $H = (\mathrm{S}_n)_x$ for a $x \in U$ and $\mathbb{R}^n_H = \{y \in \mathbb{R}^n : (\mathrm{S}_n)_y = H\}$. Since S_n is finite and the manifolds $\tau.\mathbb{R}^n_H$ for $\tau \in \mathrm{S}_n$ either coincide or are pairwise disjoint, the components of $\mathrm{S}_n.\mathbb{R}^n_H$ are open subsets of $\tau.\mathbb{R}^n_H$ for $\tau \in \mathrm{S}_n$. Thus, we may assume that S is a component of $\mathbb{R}^n_H \cap U$.

Denote by π the quotient projection $N(U) \to N(U)/Z(U) = W$. For any $u \in U$ we have $W_u = \pi(N(U) \cap (\mathrm{S}_n)_u)$ and thus $\mathbb{R}^n_H \cap U \subseteq \{u \in U : W_u = W_x\}$. By definition and a similar argument as above, the components of the subset $\{u \in U : W_u = W_x\}$ are orbit type strata of U. So the ambient stratum S is contained in a unique orbit type stratum R_S.

To (2): Let R be an orbit type stratum and let $\mathfrak{S}$ be the set of all ambient strata S such that $R_S = R$, where R_S is the unique orbit type stratum from (1). Clearly, $R = \bigcup \mathfrak{S}$ and for each $S \in \mathfrak{S}$ we have $\dim S \leq \dim R$. Since the set $\mathfrak{S}$ is finite, there is a stratum $S \in \mathfrak{S}$ such that $\dim S = \dim R$. □

Remarks. (1) It is easy to see that proposition 13.3.2 is true if one replaces the S_n-module $\mathbb{R}^n$ by any finite dimensional G-module V, where G is a finite group.

(2) Proposition 13.3.2 implies that the orbit type stratification of U is coarser than its ambient stratification. That means, by definition 9.6.4, that for each

ambient stratum S there exists an orbit type stratum R_S such that $S \subseteq R_S$, $\mathrm{id}|_S : S \to R_S$ is smooth, and for all $S \subseteq \overline{S'}$ we have $R_S \subseteq \overline{R_{S'}}$. It remains to check the last condition: Assume that $S \subseteq \overline{S'}$. Since $S \subseteq R_S$ and $S \subseteq \overline{S'} \subseteq \overline{R_{S'}}$, we obtain $R_S \cap \overline{R_{S'}} \neq \emptyset$, and, by the frontier condition, $R_S \subseteq \overline{R_{S'}}$.

13.3.3. Orbit type and ambient stratification of U/W. Assume that the restrictions $E_i|_U$, $1 \le i \le n$, generate the algebra $\mathbb{R}[U]^W$. It follows that $E|_U = (E_1|_U, \dots, E_n|_U)$ induces a homeomorphism between U/W and the semialgebraic subset $E(U)$ of $\mathbb{R}^n/\mathrm{S}_n = E(\mathbb{R}^n) = \mathrm{Hyp}_n$. We have seen in section 9.6 that $U_{(H)} \to U_{(H)}/W$, where $H = W_u$ for some $u \in U$, is a Riemannian submersion. Since W is finite, it is even a local diffeomorphism. By proposition 13.3.2, this implies that for any ambient stratum S in U the image $E(S)$ is a smooth manifold. The collection $\mathcal{T} = \{E(S) : S \text{ ambient stratum in } U\}$ obviously coincides with the collection obtained by restricting to $E(U)$ the orbit type stratification of $\mathbb{R}^n/\mathrm{S}_n = E(\mathbb{R}^n) = \mathrm{Hyp}_n$. It is easily verified that the frontier condition for the orbit type stratification of $\mathbb{R}^n/\mathrm{S}_n = E(\mathbb{R}^n) = \mathrm{Hyp}_n$ implies the frontier condition for $\mathcal{T}$. Consequently, $\mathcal{T}$ provides a stratification of $E(U)$. Let us denote this stratification as the *ambient stratification of* $E(U)$.

13.3.4. Normal nonflatness with respect to ambient and orbit type stratification. Consider a smooth curve $c : \mathbb{R} \to E(U) = U/W$ (smooth as curve in $\mathbb{R}^n$). It may then be also viewed as a smooth curve in $\mathbb{R}^n/\mathrm{S}_n = E(\mathbb{R}^n) = \mathrm{Hyp}_n$. Thus it makes sense to speak about the normal nonflatness of c at some point t_0 with respect to the orbit type stratification of U/W on the one hand and with respect to the orbit type stratification of $\mathbb{R}^n/\mathrm{S}_n$ on the other hand. To shorten notation we shall say that c is normally nonflat at t_0 with respect to the ambient stratification of U/W if and only if it is normally nonflat at t_0 with respect to the orbit type stratification of $\mathbb{R}^n/\mathrm{S}_n$.

PROPOSITION. *Let U be a linear subspace in $\mathbb{R}^n$ endowed with the induced action by $W = W(U)$ and assume that the restrictions $E_i|_U$, $1 \le i \le n$, generate $\mathbb{R}[U]^W$. Consider a smooth curve $c : \mathbb{R} \to E(U) = U/W$. If c is normally nonflat at t_0 with respect to the ambient stratification of U/W, then it is normally nonflat at t_0 with respect to the orbit type stratification of U/W.*

PROOF. The set of reflection hyperplanes H of the reflection group S_n is in bijective correspondence with the set of linear functionals ω_H on $\mathbb{R}^n$ of the form $x_j - x_i$ for $1 \le i < j \le n$, namely H is the kernel of ω_H. Let us consider the restrictions $\omega_H|_U$ to U. If c is normally nonflat at t_0 with respect to the ambient stratification, then, by lemma 13.1.1, any two of the increasingly ordered continuous roots of the polynomial $c(t) \in E(U) \subseteq \mathrm{Hyp}_n$ either coincide identically near t_0 or do not meet at t_0 of infinite order. Then for the continuous lift $\bar{c}$ of c defined by such a choice of roots any function $\omega_H \circ \bar{c}$ either vanishes identically near t_0 or does not vanish at t_0 of infinite order.

Let s be a minimal integer such that $c(t)$ lies in $A_{s,\mathrm{orb}}$ for t near t_0, where $A_{s,\mathrm{orb}}$ is the union of all orbit type strata of U/W of dimension $\le s$.

Denote by π_U the projection $U \to U/W$. Let R be an orbit type stratum contained in $\pi_U^{-1}(A_{s-1,\mathrm{orb}})$ and let $S_1, \dots, S_k$ be the ambient strata of the same dimension as R contained in R (see proposition 13.3.2). For each $1 \le j \le k$ denote by $\mathcal{H}_j$ the set of reflection hyperplanes for reflections in S_n fixing S_j pointwise. Let Ω_j be the set of linear functionals $\omega_H|_U$ for $H \in \mathcal{H}_j$. Put $f_{R,j} = \sum_{\omega \in \Omega_j} \omega^2$. By definition, the equation $f_{R,j} = 0$ defines a linear subspace of U in which S_j is an open subset. Let $f_R = \prod_{j=1}^k f_{R,j}$. Consider the natural action of W on $\mathbb{R}[U]$ and let $W.f_R = \{f_R^1, \dots, f_R^l\}$ be the orbit through f_R with respect to this action.

Define $F_R = f_R^1 \cdots f_R^l$. By construction, $F_R \in \mathbb{R}[U]^W$, and the set Z_R of zeros of F_R viewed as a function on U/W is contained in $A_{s-1,\mathrm{orb}}$. Moreover, $A_{s-1,\mathrm{orb}}$ is the union of the Z_R, where R ranges over all orbit type strata (of maximal dimension) contained in $\pi_U^{-1}(A_{s-1,\mathrm{orb}})$. Thus $F = \prod_R F_R$, where the product is taken over all orbit type strata (of maximal dimension) R contained in $\pi_U^{-1}(A_{s-1,\mathrm{orb}})$, is a regular function on U/W whose set of zeros equals $A_{s-1,\mathrm{orb}}$. By construction, the function $F \circ c$ is nonflat at t_0.

This proves the statement. □

We define $F_{\mathrm{amb}}(c)$ (resp. $F_{\mathrm{orb}}(c)$) to be the set of all $t \in \mathbb{R}$ such that c is normally flat at t with respect to the ambient (resp. orbit type) stratification of $E(U)$. It follows that in the situation of proposition 13.3.4 we have $F_{\mathrm{orb}}(c) \subseteq F_{\mathrm{amb}}(c)$.

13.4. Choosing roots of polynomials with symmetries smoothly

13.4.1. Consider a smooth curve of hyperbolic polynomials

$$P(t)(x) = x^n - a_1(t)x^{n-1} + a_2(t)x^{n-2} - \cdots + (-1)^n a_n(t) \qquad (t \in \mathbb{R}).$$

We are interested in conditions that guarantee the existence of a smooth parameterization of the roots of P. Such conditions have been found in theorem 4.3.1. There no additional assumptions on the polynomials $P(t)$ have been made.

In this section we are going to improve those results if the set of roots $x_1(t), \ldots, x_n(t)$ of $P(t)$ has symmetries additional to its invariance under permutations.

Let as assume that the additional symmetries of $P(t)$ are given by linear relations between the roots of $P(t)$. Otherwise put, there is a linear subspace U of $\mathbb{R}^n$ such that $(x_1(t), \ldots, x_n(t)) \in U$ for all $t \in \mathbb{R}$. Then, the curve $P(t)$ lies in the semialgebraic subset $E(U)$ of $\mathrm{Hyp}_n = E(\mathbb{R}^n) = \mathbb{R}^n/\mathrm{S}_n$, the space of hyperbolic polynomials of degree n.

The linear subspace $U \subseteq \mathbb{R}^n$ inherits an effective action by the group $W = W(U)$.

Let us suppose that the restrictions $E_i|_U$, $1 \le i \le n$, generate the algebra $\mathbb{R}[U]^W$. Then $E|_U = (E_1|_U, \ldots, E_n|_U)$ induces a homeomorphism between U/W and the semialgebraic subset $E(U)$ of Hyp_n.

Lemma. *Consider a continuous curve of hyperbolic polynomials*

$$P(t)(x) = x^n - a_1(t)x^{n-1} + a_2(t)x^{n-2} - \cdots + (-1)^n a_n(t) \qquad (t \in \mathbb{R}).$$

Let U be some linear subspace of $\mathbb{R}^n$ and assume that the restrictions $E_i|_U$, $1 \le i \le n$, generate the algebra $\mathbb{R}[U]^{W(U)}$. Then the following two conditions are equivalent:

(1) *There exists a continuous parameterization $x(t)$ of the roots $x_1(t), \ldots, x_n(t)$ of $P(t)$ such that $x(t) \in U$ for all $t \in \mathbb{R}$.*
(2) *$P(t) \in E(U)$ for all $t \in \mathbb{R}$.*

Proof. The implication (1) ⇒ (2) is trivial. Suppose that $P(t)$ is a continuous curve in $E(U)$. By assumption, we may view $P(t)$ as a curve in the orbit space $U/W(U) \cong E(U)$. It allows a continuous lift $x(t)$ into U, by theorem 11.1.1, which constitutes a parameterization of the roots of $P(t)$. □

13.4.2. The smooth curve of polynomials $P(t)$ which lies in $E(U)$ may be viewed as a smooth curve in the orbit space U/W. A smooth lift of $P(t)$ over the orbit map $E|_U$ to the W-module U provides a smooth parameterization of the roots of the polynomials $P(t)$.

By theorem 10.2.1, we may conclude: If $P(t)$ is normally nonflat at $t = t_0$ with respect to the orbit type stratification of $E(U)$, then $P(t)$ is smoothly solvable near $t = t_0$.

Consider the closed sets $F_{\text{amb}}(P)$ and $F_{\text{orb}}(P)$, as defined in section 13.3. By proposition 13.3.4, the set $F_{\text{orb}}(P)$ is contained in $F_{\text{amb}}(P)$. We have found that that $P(t)$ is smoothly solvable locally near any $t_0 \in \mathbb{R}\backslash F_{\text{orb}}(P)$. Any two smooth parameterizations of the roots of $P(t)$ near such a t_0 differ by a constant permutation, see theorem 4.3.1. Thus the local solutions may be glued to a smooth solution on $\mathbb{R}\backslash F_{\text{orb}}(P)$.

It follows from theorem 7.1.1 that any smooth curve of monic hyperbolic polynomials of fixed degree allows a global twice differentiable parameterization of its roots. By the methods used in the proof of theorem 7.1.1, it is easy to combine this with the result above in order to get the following theorem.

THEOREM. *Consider a smooth curve of hyperbolic polynomials*

$$P(t)(x) = x^n - a_1(t)x^{n-1} + a_2(t)x^{n-2} - \cdots + (-1)^n a_n(t) \qquad (t \in \mathbb{R}).$$

Let U be some linear subspace of $\mathbb{R}^n$ such that:

(1) *The restrictions $E_i|_U$, $1 \le i \le n$, generate the algebra $\mathbb{R}[U]^{W(U)}$.*
(2) *$P(t) \in E(U)$ for all $t \in \mathbb{R}$.*

Then: There exists a global twice differentiable parameterization of the roots of $P(t)$ on $\mathbb{R}$ which is smooth on $\mathbb{R}\backslash F_{\text{orb}}(P)$. □

REMARK. The orbit type stratification and the ambient stratification of $E(U)$ do in general not coincide, whence theorem 13.4.2 provides an actual improvement of the statement of theorem 4.3.1. In other words, in general we have $F_{\text{orb}}(P) \subsetneq F_{\text{amb}}(P)$. It may, for instance, happen that $P(0)$ is regular in $E(U) = U/W$ but singular in $\text{Hyp}_n = \mathbb{R}^n/\,\text{S}_n$ and $P(t)$ is normally flat at $t = 0$ with respect to the ambient stratification. See examples in section 13.7.

13.4.3. Let us suppose that a linear subspace U of $\mathbb{R}^n$ is given. It is then a purely computational problem to check whether the assumptions we have made in the forgoing discussion are satisfied. There are algorithms in computational invariant theory (e.g. [**DK02**], [**Stu93**]) which allow to decide whether the restrictions $E_i|_U$, $1 \le i \le n$, generate the algebra $\mathbb{R}[U]^{W(U)}$. If the answer is yes, theorem 9.2.2 provides an explicit way to describe the semialgebraic subset $E(U) \subseteq \text{Hyp}_n$ by a finite set of polynomial equations and inequalities. So the condition that the curve P lies in $E(U)$ may again be check computationally. The orbit type stratification and the ambient stratification of $E(U)$ can be determined explicitly, using that the k-dimensional primary strata of $E(U)$ are the connected components of the set $\{z \in E(U) : \text{rank}\,\tilde{B}(z) = k\}$, where B is the corresponding generalized Bezoutiant defined in (9.10). Then all ingredients are supplied in order to decide whether the curve $P(t)$ is normally nonflat at some $t = t_0$ with respect to the one or the other stratification of $E(U)$.

Note that there are refined approaches and algorithms for computing the orbit space V/G and its orbit type stratification of a G-module V (when identified with the image of its orbit map). In [**SV03**] rational parameterizations of the strata are obtained, while [**Bay04**] provides an algorithm yielding a description of each stratum in terms of a minimal number of polynomial equations and inequalities, if G is finite.

We shall carry out that procedure explicitly in example 13.7.7.

13.5. Choosing roots of polynomials with symmetries differentiably

Consider a curve of hyperbolic polynomials

$$P(t)(x) = x^n - a_1(t)x^{n-1} + a_2(t)x^{n-2} - \cdots + (-1)^n a_n(t) \qquad (t \in \mathbb{R}).$$

We have proved the following results:

13.5.1. Result. We have:

(1) If all a_i are of class C^n, then there exists a differentiable parameterization of the roots of $P(t)$ with locally bounded derivative, theorem 5.5.13 or theorem 6.3.1.
(2) If all a_i are of class C^{2n}, then any differentiable parameterization of the roots of $P(t)$ is actually C^1, theorem 7.1.1.
(3) If all a_i are of class C^{3n}, then there exists a twice differentiable parameterization of the roots of $P(t)$, theorem 7.1.1.

In chapter 12 we have proved the following generalizations:

13.5.2. Result. Let $\rho : G \to \mathrm{O}(V)$ be a finite dimensional representation of a finite group G. Let $d = d(\rho)$ be the maximum of the degrees of a minimal system of homogeneous generators $\sigma_1, \dots, \sigma_m$ of $\mathbb{R}[V]^G$. Write $V = V_1 \oplus \cdots \oplus V_l$ as orthogonal direct sum of irreducible subspaces V_i. Define $k_i := \min\{|G.v| : v \in V_i \setminus \{0\}\}$, $1 \le i \le l$, and $k := \max\{d(\rho), k_1, \dots, k_l\}$. Let $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^m$ be a curve in the orbit space. Then:

(1) If c is of class C^k, then there exists a differentiable lift of c to V with locally bounded derivative.
(2) If c is of class C^{k+d}, then any differentiable lift of c is actually of class C^1.
(3) If c is of class C^{k+2d}, then there exists a twice differentiable lift of c to V.

Again we may use these facts in order to improve the results for curves $P(t)$ of hyperbolic polynomials with symmetries.

Let U be some linear subspace of $\mathbb{R}^n$ such that the restrictions $E_i|_U$, $1 \le i \le n$, generate the algebra $\mathbb{R}[U]^{W(U)}$, and $P(t) \in E(U)$ for all $t \in \mathbb{R}$. It follows that we may view $P(t)$ as a curve in the orbit space $U/W(U) = E(U)$, and any lift of $P(t)$ over the orbit map $E|_U$ to U gives a parameterization of the roots of $P(t)$ of the same regularity.

Provided that the integer k, associated to the $W(U)$-module U as above, is less than the degree n of the polynomials in $P(t)$, we are able, using 13.5.2, to lower the degree of regularity in the assumptions of the statements in 13.5.1. We shall give examples in section 13.7.

13.6. Construction of a class of examples

We will present a class of examples which our considerations apply to.

13.6.1. The maps $F_{G,v}$ and $P_{G,v}$. Let $G \subseteq \mathrm{O}(V)$ be a finite group whose action on the vector space V is irreducible and effective. Choose some non-zero orbit $G.v$. Introducing some numbering we can write $G.v = \{g_1.v, \dots, g_n.v\}$, where $|G.v| = n$ and $g_i \in G$. We define a mapping $F_{G,v} : V \to \mathbb{R}^n$ by

$$F_{G,v}(x) := (\langle g_1.v \mid x\rangle, \dots, \langle g_n.v \mid x\rangle).$$

Since the linear span of $G.v$ spans V, the mapping $F_{G,v}$ is a linear isomorphism onto its image $F_{G,v}(V) =: U_{G,v}$. The linear space $U_{G,v} \subseteq \mathbb{R}^n$ carries the action of $W_{G,v} := W(U_{G,v})$ and a natural G-action given by transformations from $W_{G,v}$. Since the G-action is irreducible, so is the $W_{G,v}$-action. Hence $U_{G,v} \subseteq \{y \in \mathbb{R}^n : y_1 + \cdots + y_n = 0\}$. Irreducibility and effectiveness of the G-action induce an injection

$G \hookrightarrow W_{G,v}$. Thus we may consider G as a subgroup of $W_{G,v}$, and in this picture $F_{G,v}$ is G-equivariant.

REMARK. The linear space $U_{G,v}$ always intersects the submanifold of regular points in the S_n-module $\mathbb{R}^n$. Namely: For $1 \le i < j \le n$ we define $U_{i,j} = \{F_{G,v}(x) : \langle g_i.v \mid x\rangle = \langle g_j.v \mid x\rangle, x \in V\}$. By definition, $U_{i,j}$ is a linear subspace of $U_{G,v}$ and $\bigcup_{i<j} U_{i,j}$ is the set of singular points of the S_n-module $\mathbb{R}^n$ contained in $U_{G,v}$. Since, by definition, $g_i.v \neq g_j.v$ for any $i < j$, we have $\dim U_{i,j} = n-1$. Thus, $\bigcup_{i<j} U_{i,j} \neq U_{G,v}$, which gives the assertion.

Put $P_{G,v} := E \circ F_{G,v}$. Then $P_{G,v}$ is proper, since E and $F_{G,v}$ are proper.

LEMMA. *Suppose that $P_{G,v}$ separates G-orbits. Then we have $G = W_{G,v}$.*

PROOF. The groups G and $W_{G,v}$ have the same orbits in $U_{G,v}$. For: Suppose that $\tau \in W_{G,v}$ and $x, y \in V$ such that $F_{G,v}(y) = \tau.F_{G,v}(x)$. Since $P_{G,v}$ separates orbits, it follows that there exists some $g \in G$ such that $y = g.x$, whence $g.F_{G,v}(x) = \tau.F_{G,v}(x)$.

Now choose $x \in V$ such that $F_{G,v}(x)$ is a regular point of the $W_{G,v}$-module $U_{G,v}$. The regular points of any effective linear finite group representation are precisely those with trivial isotropy groups. We may conclude that x is a regular point of the G-module V. So $|W_{G,v}| = |W_{G,v}.F_{G,v}(x)| = |G.x| = |G|$, and thus $G = W_{G,v}$. □

If $P_{G,v}$ separates G-orbits, then, by the above lemma, the $G = W_{G,v}$-modules V and $U_{G,v}$ are equivalent. In particular, it follows that the restriction $E|_{U_{G,v}}$ separates $W_{G,v}$-orbits, $F_{G,v}$ induces a homeomorphism between V/G and $U_{G,v}/W_{\rho,v}$, and $F^*_{G,v} : \mathbb{R}[U_{G,v}]^{W_{G,v}} \to \mathbb{R}[V]^G$ is an algebra isomorphism.

PROPOSITION. *The following conditions are equivalent:*

(1) *$P_{G,v}$ separates G-orbits.*
(2) *For all $x \in V$ we have $F_{G,v}(G.x) = \mathrm{S}_n .F_{G,v}(x) \cap U_{G,v}$.*
(3) *$P_{G,v}$ induces a homeomorphism between V/G and $P_{G,v}(V)$.*

PROOF. Since E separates S_n-orbits, for each $x \in V$ there exists a $z \in \mathbb{R}^n$ such that $E^{-1}(z) = \mathrm{S}_n .F_{G,v}(x)$. Then the equivalence of (1) and (2) follows from

$$P^{-1}_{G,v}(z) = F^{-1}_{G,v}(\mathrm{S}_n .F_{G,v}(x)) = F^{-1}_{G,v}(\mathrm{S}_n .F_{G,v}(x) \cap U_{G,v}).$$

The equivalence of (1) and (3) follows easily from lemma 8.2.3. □

Note that the introduced construction of $F_{G,v}$ and $P_{G,v}$ essentially coincides with the construction of orbit Chern classes as described in 13.2.2.

13.6.2. Uniqueness. Let us discuss uniqueness of the above construction. Suppose $G \subseteq \mathrm{O}(V)$ is a finite group. Denote by $\mathrm{Aut}(G)$ the group of automorphisms of G. Let S be the set of all reflections belonging to G. Denote by $\mathrm{Aut}(G, S)$ the group of automorphisms of G preserving the set S. Let $a \in \mathrm{Aut}(G, S)$. A diffeomorphism $T : V \to V$ is called *a-equivariant*, if $T \circ g = a(g) \circ T$ for any $g \in G$ (cf. [**Los01**]).

LEMMA. *Suppose $G \subseteq \mathrm{O}(V)$ is a finite group. Let $a \in \mathrm{Aut}(G, S)$ and let $T : V \to V$ be an a-equivariant diffeomorphism. Then the isotropy groups of x and $T(x)$ are isomorphic, for all $x \in V$, T maps orbits onto orbits, and T induces an automorphism of the orbit type stratification of V.*

PROOF. It is easily seen that $G_{T(x)} = a(G_x)$ and $T(G.x) = G.T(x)$ for all $x \in V$. Further, it is evident that $G_x = gHg^{-1}$ if and only if $G_{T(x)} = a(g)a(H)a(g)^{-1}$. The statement follows. □

Let $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a smooth curve and $\bar{c} : \mathbb{R} \to V$ a smooth lift of c. The orbit space V/G has a smooth structure given by the sheaf $C^\infty(V/G) = C^\infty(V)^G$ of smooth G-invariant functions on V. Then c induces a continuous algebra morphism $c^* : C^\infty(V/G) \to C^\infty(\mathbb{R})$ and $\bar{c}$ induces a continuous algebra morphism $\bar{c}^* : C^\infty(V) \to C^\infty(\mathbb{R})$ such that $c^* = \bar{c}^* \circ \sigma^*$. This algebraic lifting problem is equivalent to the geometrical one. It is evident that to determine $\bar{c}^*$ it suffices to know the images under $\bar{c}^*$ of some system of global coordinate functions $x_1, \dots, x_m$, where $m = \dim V$. The same is true for c^*, and in this case we may take the basic invariants $\sigma_1, \dots, \sigma_n$ as global coordinates functions, by Schwarz's theorem 8.2.5. If $f : V/G \to V/G$ is a smooth diffeomorphism one can take instead of the σ_i the functions $f^*(\sigma_i)$ with the same result. Thus, the problem of smooth lifting is invariant with respect to the group of diffeomorphisms of V/G. Each such diffeomorphism has a smooth lift to V which is an a-equivariant diffeomorphism, for some $a \in \mathrm{Aut}(G, S)$, see [**Los01**] or theorem 16.1.6. Conversely, any smooth a-equivariant diffeomorphism of V induces a smooth diffeomorphism of V/G, by the above lemma.

Therefore, we may regard two constructions as described above, carried out for distinct points v and w in V, as equivalent with respect to our lifting problem, if there exists a smooth a-equivariant diffeomorphism $T : V \to V$ with $v = T(w)$, for some $a \in \mathrm{Aut}(G, S)$.

If T is of a particular form, we can even say more.

Proposition. *Suppose $G \subseteq \mathrm{O}(V)$ is a finite group. Let $v, w \in V \backslash \{0\}$. If there exists a homothety or an a-equivariant linear orthogonal map $T : V \to V$, for some $a \in \mathrm{Aut}(G, S)$, such that $v = T(w)$, then $P_{G,v}(V)$ and $P_{G,w}(V)$ are homeomorphic, and $\mathbb{R}[E_1 \circ F_{G,v}, \dots, E_n \circ F_{G,v}]$ and $\mathbb{R}[E_1 \circ F_{G,w}, \dots, E_n \circ F_{G,w}]$ are isomorphic.*

Moreover, in both cases, the ambient stratifications of $U_{G,v}$ and $U_{G,w}$ are isomorphic, i.e., there exists a linear isomorphism $U_{G,v} \to U_{G,w}$ mapping strata onto strata.

Proof. If T is a homothety, then it is equivariant ($a = \mathrm{id}$) and $U_{G,v} = U_{G,w}$. If T is a-equivariant linear orthogonal, then, by the above lemma, the linear subspaces $U_{G,v}$ and $U_{G,w}$ of $\mathbb{R}^n$ differ only by a permutation from S_n. In both cases $P_{G,v}(V)$ and $P_{G,w}(V)$ are homeomorphic, and $T^* : \mathbb{R}[E_1 \circ F_{G,v}, \dots, E_n \circ F_{G,v}] \to \mathbb{R}[E_1 \circ F_{G,w}, \dots, E_n \circ F_{G,w}]$ is an algebra isomorphism.

The supplement in the lemma follows immediately from the fact that $U_{G,v}$ and $U_{G,w}$ differ only by a permutation of S_n. □

13.6.3. If $P(t)$ is a smooth curve of hyperbolic polynomials lying in $P_{G,v}(V)$ and provided that the polynomials $E_i \circ F_{G,v}$, $1 \le i \le n$, generate $\mathbb{R}[V]^G$, we may apply the results of sections 13.4 and 13.5.

We will investigate the case of finite reflection groups in the next section.

13.7. Finite reflection groups

Suppose U is a linear subspace of $\mathbb{R}^n$. Let the symmetric group S_n act on $\mathbb{R}^n$ by permuting the coordinates and endow U with the induced action of $W = W(U)$. We shall assume in this section that W is a finite reflection group.

13.7.1. Remark. If W is a finite reflection group, proposition 13.3.2 reduces to the following statement: *Any reflection hyperplane of W in U is the intersection with U of some reflection hyperplane of S_n in $\mathbb{R}^n$.* For: Let H be a reflection hyperplane of W in U. By proposition 13.3.2, there exists a ambient stratum S of U such that $S \subseteq H$ and $\dim S = \dim H$. Obviously, $S \subseteq (\mathbb{R}^n)_{\mathrm{sing}} \cap U$, and so there are reflection hyperplanes $P_1, \dots, P_l$ of S_n in $\mathbb{R}^n$ which contain S. Since

$\dim S = \dim U - 1$, there is a $1 \le i \le n$ such that $P_i \cap U$ is a hyperplane in U. Since S is contained in both H and $P_i \cap U$, we have $H = P_i \cap U$.

13.7.2. For any finite reflection group $W \subseteq \mathrm{O}(U)$ we may write U as the orthogonal direct sum of W-invariant subspaces $U_0 = U^W, U_1, \dots, U_l$ such that W is isomorphic to $W_0 \times W_1 \times \cdots \times W_l$, where $W_i = \{\tau|_{U_i} : \tau \in W\}$. Each W_i $(i \ge 1)$ is one of the groups (compare with theorem 12.7.3)

$$\mathrm{A}_m, m \ge 1; \mathrm{B}_m, m \ge 2; \mathrm{D}_m, m \ge 4; \mathrm{I}_2^m, m \ge 5, m \ne 6;$$
$$\mathrm{G}_2; \mathrm{H}_3; \mathrm{H}_4; \mathrm{F}_4; \mathrm{E}_6; \mathrm{E}_7; \mathrm{E}_8\,.$$

It follows that $\mathbb{R}[U]^W \cong \mathbb{R}[U_1]^{W_1} \otimes \cdots \otimes \mathbb{R}[U_l]^{W_l}$ and $U/W \cong U_1/W_1 \times \cdots \times U_l/W_l$. A smooth curve $c = (c_1, \dots, c_l)$ in the orbit space U/W is then smoothly liftable to U if and only if, for all $1 \le i \le l$, c_i is smoothly liftable to U_i. Note that the orbit type stratification of U/W coincides with the product stratification of the orbit type stratifications $\mathcal{Z}_i$ of the factors U_i/W_i, i.e., the strata of U/W are $S_1 \times \cdots \times S_l$, where $S_i \in \mathcal{Z}_i$. Consequently, in order to apply the results of section 13.4 and section 13.5 we may consider each factor U_i/W_i separately. So let us assume that U is an irreducible W-module.

13.7.3. To this end we have to check whether the restrictions $E_i|_U$, $1 \le i \le n$, generate the algebra $\mathbb{R}[U]^W$. In practice this is easily accomplishable: The unique degrees $d_1, \dots, d_m$, where $m = \dim U$, of the elements in a minimal system of homogeneous generators of $\mathbb{R}[U]^W$ are well known, see 13.8. It suffices to compute the Jacobian J of the polynomials $E_{d_i}|_U$, $1 \le i \le m$. If $J \ne 0 \in \mathbb{R}[U]$ then they generate $\mathbb{R}[U]^W$, see the proposition and the theorem below. Note that a necessary condition for the $E_i|_U$, $1 \le i \le n$, to generate $\mathbb{R}[U]^W$ is that the degrees $d_1, \dots, d_m$ must be pairwise distinct, see remark 13.7.4.

Proposition. *The set $\{f_1, \dots, f_n\} \subseteq \mathbb{R}[x] = \mathbb{R}[x_1, \dots, x_n]$ is algebraically independent if and only if for the Jacobian $J = \det(\frac{\partial f_i}{\partial x_j})_{ij}$ we have $J \ne 0 \in \mathbb{R}[x]$.*

Proof. For contradiction, let $0 \ne h \in \mathbb{R}[x]$ of minimal degree such that $h(f_1, \dots, f_n) = 0$. Differentiation using the chain rule yields a system of linear equations over the field of fractions $\mathbb{R}(x)$:

$$\left(\frac{\partial h}{\partial x_i}(f_1, \dots, f_n)\right)_{1 \le i \le n} \cdot \left(\frac{\partial f_i}{\partial x_j}\right)_{1 \le i,j \le n} = 0.$$

There is a $1 \le i \le n$ such that $\frac{\partial h}{\partial x_i}$ does not vanish, since h is not constant, and, by the choice of h, also $\frac{\partial h}{\partial x_i}(f_1, \dots, f_n) \ne 0 \in \mathbb{R}[x]$. Thus the above system of linear equations has a non-trivial solution, forcing $J = 0 \in \mathbb{R}[x]$.

Let, conversely, $\{f_1, \dots, f_n\}$ be algebraically independent. As the transcendence degree of $\mathbb{R}(x)$ over $\mathbb{R}$ is n (e.g. [**Lan02**, 8.1]), the sets $\{x_k, f_1, \dots, f_n\}$ are algebraically dependent for each fixed $1 \le k \le n$. Let $0 \ne h_k \in \mathbb{R}[x_0, x_1, \dots, x_n] = \mathbb{R}[x^+]$ of minimal degree such that $h_k(x_k, f_1, \dots, f_n) = 0$. Differentiation yields

$$\left(\frac{\partial h_k}{\partial x_i}(x_k, f_1, \dots, f_n)\right)_{ki} \cdot \left(\frac{\partial f_i}{\partial x_j}\right)_{ij} = \operatorname{diag}\left(-\frac{\partial h_k}{\partial x_0}(x_k, f_1, \dots, f_n)\right)_k.$$

Since $\{f_1, \dots, f_n\}$ is algebraically independent, h_k must have positive degree in x_0. So $\frac{\partial h_k}{\partial x_0} \ne 0 \in \mathbb{R}[x^+]$ and of smaller degree than h_k, forcing $\frac{\partial h_k}{\partial x_0}(x_k, f_1, \dots, f_n) \ne 0 \in \mathbb{R}[x]$. Hence,

$$\det \operatorname{diag}\left(-\frac{\partial h_k}{\partial x_0}(x_k, f_1, \dots, f_n)\right)_k \ne 0 \in \mathbb{R}[x^+],$$

and, therefore, $J \ne 0 \in \mathbb{R}[x]$. □

THEOREM. *Let a finite group G act faithfully and orthogonally on a finite dimensional vector space V, and let $f_1, \dots, f_n \in \mathbb{R}[V]^G$ be homogeneous invariants with $n = \dim V$. Then the following statements are equivalent:*
(1) $\mathbb{R}[V]^G = \mathbb{R}[f_1, \dots, f_n]$.
(2) *The f_i are algebraically independent and* $\prod_{i=1}^n \deg f_i = |G|$.

PROOF. To (2) $\Rightarrow$ (1): Let H be the subgroup of G generated by $S(G)$, the set of reflections in G. By Shephard and Todd's theorem (e.g. [**DK02**], [**Smi95**]), we find $\mathbb{R}[V]^H = \mathbb{R}[h_1, \dots, h_n]$ for $h_i \in \mathbb{R}[V]^H$. Set $d_i = \deg f_i$ and $e_i = \deg h_i$, and arrange $d_1, \dots, d_n$ and $e_1, \dots, e_n$ in non-decreasing order. Then we claim that $d_i \geq e_i$ for $1 \leq i \leq n$. We have

$$\mathbb{R}[f_1, \dots, f_n] \subseteq \mathbb{R}[V]^G \subseteq \mathbb{R}[V]^H = \mathbb{R}[h_1, \dots, h_n].$$

It follows that $d_1 \geq e_1$. Assume that $d_i \geq e_i$ for $1 \leq i \leq m$ and consider f_{m+1}. If $d_{m+1} < e_{m+1}$, then f_{m+1} must be a polynomial in $h_1, \dots, h_m$. Since $d_i \leq d_{m+1} < e_{m+1}$ for $1 \leq i \leq m$, also $f_1, \dots, f_m$ are polynomials in $h_1, \dots, h_m$, and, hence, $f_1, \dots, f_{m+1} \in \mathbb{R}[h_1, \dots, h_m]$ would be algebraically independent which is impossible. The claim follows.

Both G and H have the same reflections, so, by the lemmas below,

$$\sum_{i=1}^n (d_i - 1) \leq |S(G)| = |S(H)| = \sum_{i=1}^n (e_i - 1).$$

Therefore, $d_i = e_i$ for $1 \leq i \leq n$, and the statement follows.

To (1) $\Rightarrow$ (2): This follows from the lemmas below. □

LEMMA. *Let a finite group G act faithfully and orthogonally on a finite dimensional vector space V. Then the Laurent expansion of the Poincaré series $P(\mathbb{R}[V]^G, t)$ begins as follows:*

$$P(\mathbb{R}[V]^G, t) = \frac{\frac{1}{|G|}}{(1-t)^n} + \frac{\frac{|S(G)|}{2|G|}}{(1-t)^{n-1}} + \cdots, \tag{13.2}$$

where $S(G)$ is the set of reflections in G.

PROOF. By Molien's theorem (e.g. [**DK02**], [**Smi95**]),

$$P(\mathbb{R}[V]^G, t) = \frac{1}{|G|} \sum_{g \in G} \frac{1}{\det(1 - g^{-1}t)} = \frac{1}{|G|} \left(\frac{1}{\det(1-t)} + \sum_{1 \neq g \in G} \frac{1}{\det(1 - g^{-1}t)} \right)$$

None of the group elements $g \neq 1$ can have 1 as an eigenvalue of order n. This gives the leading term of (13.2).

Now, $\sum_{1 \neq g \in G} \frac{1}{\det(1-g^{-1}t)}$ has a pole of order at most $n-1$ at $t = 1$. An element $g \in G$ contributes to this pole if and only if $\dim V^g = n - 1$, i.e., if g is a reflection. Then we may write

$$P(\mathbb{R}[V]^G, t) = \frac{1}{|G|} \left(\frac{1}{(1-t)^n} + \sum_{1 \neq g \in S(G)} \frac{1}{\det(1 - g^{-1}t)} + \cdots \right).$$

For a reflection g we have $\det(1 - g^{-1}t) = (1-t)^{n-1}(1+t)$. It follows that the contribution of g to the second term in the Laurent series of $P(\mathbb{R}[V]^G, t)$ about $t = 1$ is $\frac{1}{2|G|}$. This completes the proof. □

LEMMA. *Let a finite group G act faithfully and orthogonally on a finite dimensional vector space V with $n = \dim V$. Denote by $S(G)$ the set of reflections in G.*

(1) *If* $\mathbb{R}[V]^G = \mathbb{R}[f_1, \dots, f_n]$, *then* $|G| = d_1 \cdots d_n$ *and* $|S(G)| = \sum_{i=1}^n (d_i - 1)$, *where* $d_i = \deg f_i$.
(2) *If the* $f_1, \dots, f_n \in \mathbb{R}[V]^G$ *are algebraically independent and* $\prod_{i=1}^n d_i = |G|$, *where* $d_i = \deg f_i$, *then* $|S(G)| \geq \sum_{i=1}^n (d_i - 1)$.

PROOF. To (1): For the Poincaré series of $\mathbb{R}[V]^G$ we find, by assumption,

$$P(\mathbb{R}[V]^G, t) := \sum_{d=0}^{\infty} \dim_{\mathbb{R}} \mathbb{R}[V]^G_d t^d = \prod_{i=1}^{n} \frac{1}{1 - t^{d_i}},$$

where $d_i = \deg f_i$. Now

$$\prod_{i=1}^{n} \frac{1}{1 - t^{d_i}} = (1-t)^{-n} \prod_{i=1}^{n} \frac{1}{\sum_{j=0}^{d_i - 1} t^j}$$

and

$$\left. \prod_{i=1}^{n} \frac{1}{\sum_{j=0}^{d_j - 1} t^j} \right|_{t=1} = \prod_{i=1}^{n} \frac{1}{d_i}.$$

We expand $P(\mathbb{R}[V]^G, t)$ in a Laurent series about $t = 1$:

$$P(\mathbb{R}[V]^G, t) = \frac{\frac{1}{d_1 \cdots d_n}}{(1-t)^n} + \cdots .$$

Equating the leading coefficient with the one given in (13.2) gives $|G| = d_1 \cdots d_n$.

Next we compute the coefficient of $\frac{1}{(1-t)^{n-1}}$. Consider

$$\begin{aligned}
\frac{1}{1-t^d} - \frac{1}{d}\frac{1}{1-t} &= \frac{1}{d}\left(\frac{d}{1-t^d} - \frac{1}{1-t}\right) = \frac{1}{d}\left(\frac{d - (1 + t + \cdots + t^{d-1})}{1-t^d}\right) \\
&= \frac{1}{d}\left(\frac{1-1}{1-t^d} + \frac{1-t}{1-t^d} + \cdots + \frac{1-t^{d-1}}{1-t^d}\right) \\
&= \frac{1}{d}\left(\frac{1}{1 + t + \cdots + t^{d-1}} + \cdots + \frac{1 + t + \cdots + t^{d-2}}{1 + t + \cdots + t^{d-1}}\right),
\end{aligned}$$

which by evaluating at $t = 1$ gives

$$\left[\frac{1}{1-t^d} - \frac{1}{d}\frac{1}{1-t}\right]_{t=1} = \frac{d-1}{2d}.$$

It follows that (e.g. by formal differentiation)

$$P(\mathbb{R}[V]^G, t) = \frac{\frac{1}{d_1 \cdots d_n}}{(1-t)^n} + \frac{\frac{1}{2 d_1 \cdots d_n} \sum_{i=1}^n (d_i - 1)}{(1-t)^{n-1}} + \cdots .$$

Equating this expression with (13.2) provides the statement.

To (2): By (13.2), we have

$$P(\mathbb{R}[V]^G, t) = \frac{\frac{1}{|G|}}{(1-t)^n} + \frac{\frac{|S(G)|}{2|G|}}{(1-t)^{n-1}} + \cdots ,$$

and, by (1),

$$P(\mathbb{R}[f_1, \dots, f_n], t) = \frac{\frac{1}{d_1 \cdots d_n}}{(1-t)^n} + \frac{\frac{1}{2 d_1 \cdots d_n} \sum_{i=1}^n (d_i - 1)}{(1-t)^{n-1}} + \cdots .$$

Let us treat t as a complex variable. Both $P(\mathbb{R}[V]^G, t)$ and $P(\mathbb{R}[f_1, \dots, f_n], t)$ have a pole of order n at $t = 1$, and, by assumption, we find

$$\lim_{t \to 1} (1-t)^n P(\mathbb{R}[V]^G, t) = \frac{1}{|G|} = \frac{1}{d_1 \cdots d_n} = \lim_{t \to 1} (1-t)^n P(\mathbb{R}[f_1, \dots, f_n], t).$$

Therefore, the difference $\Delta(t) = P(\mathbb{R}[V]^G, t) - P(\mathbb{R}[f_1, \dots, f_n], t)$ has a pole of order $n-1$ at $t = 1$. Since $\mathbb{R}[f_1, \dots, f_n] \subseteq \mathbb{R}[V]^G$, the Poincaré series of $\mathbb{R}[V]^G$ dominates coefficient for coefficient the Poincaré series of $\mathbb{R}[f_1, \dots, f_n]$. Hence $\Delta(t)$ is a power series with non-negative integral coefficients, and, therefore, takes on non-negative real values for t real and near 1. The same is true for the function $(1-t)^{n-1}\Delta(t)$ which is holomorphic at $t = 1$. Evaluating at $t = 1$ gives

$$0 \le \lim_{t \to 1} (1-t)^{n-1}\Delta(t) = \frac{|S(G)|}{2|G|} - \frac{\sum_{i=1}^n (d_i - 1)}{2 d_1 \cdots d_n} = \frac{1}{2|G|}\left(|S(G)| - \sum_{i=1}^n (d_i - 1)\right),$$

as desired. □

Remark. Note that statements in the proposition, the theorem, and the lemmas above are valid under more general conditions as well (cf. [**DK02**], [**Smi95**]).

13.7.4. $P_{G,v}$ for irreducible Coxeter groups G. Let us carry out the construction presented in section 13.6 for finite irreducible reflection groups $G \subseteq \mathrm{O}(V)$. Let $v \in V \backslash \{0\}$. If the polynomials $E_i \circ F_{G,v}$ generate the algebra $\mathbb{R}[V]^G$, then $W_{G,v}$ is a finite irreducible reflection group as well, by lemma 13.6.1.

Fix a system Π of simple roots of G. For any v in $C = \{x \in V : \langle x \mid r\rangle \ge 0 \text{ for all } r \in \Pi\}$, the closure of the fundamental domain associated to Π, the isotropy group G_v is generated by the simple reflections it contains, by 12.7.4.

Lemma. *Let $G \subseteq \mathrm{O}(V)$ be a finite reflection group. Each automorphism of the corresponding Coxeter diagram $\Gamma(G)$ induces an a-equivariant orthogonal automorphism of V for some $a \in \mathrm{Aut}(G, S)$.*

Proof. ([**Los01**]) Since the vertices in the Coxeter diagram $\Gamma(G)$ represent the simple roots of G, an automorphism φ of $\Gamma(G)$, defines uniquely an automorphism $a_\varphi \in \mathrm{Aut}(G, S)$. Suppose the simple roots have unit length. Since they form a basis for V the automorphism φ defines naturally an orthogonal automorphism T_φ of V. It is easily checked that T_φ is a_φ-equivariant. □

Theorem. *Suppose $G \subseteq \mathrm{O}(V)$ is a finite irreducible reflection group. Let $v \in V \backslash \{0\}$ such that the cardinality of G_v is maximal. Then: The polynomials $E_i \circ F_{G,v}$, $1 \le i \le n$, generate $\mathbb{R}[V]^G$ and $P_{G,v}$ induces a homeomorphism between V/G and $P_{G,v}(V)$ if and only if $G \ne \mathrm{D}_m$, $m \ge 4$.*

Proof. By proposition 13.6.2 and the previous lemma it suffices to check the statement for one single $v \ne 0$ with maximal G_v. Choosing $e_1 + \cdots + e_m - m e_{m+1}$, e_1, and e_1 for A_m, B_m, and I_2^m, respectively, one obtains the usual systems of basic invariants. The choice e_1 for D_m yields $F_{\mathrm{D}_m, e_1} = F_{\mathrm{B}_m, e_1}$, whence the polynomials $E_i \circ F_{\mathrm{D}_m, e_1}$, $1 \le i \le n = 2m$, cannot separate D_m-orbits. For the remaining irreducible reflection groups the necessary computations have been carried out by Mehta [**Meh88**], see also 13.8. □

Remark. If for D_m with m odd one chooses $v = e_1 + \cdots + e_m$, then the polynomials $E_i \circ F_{\mathrm{D}_m, v}$, $1 \le i \le n = 2^{m-1}$, generate $\mathbb{R}[\mathbb{R}^m]^{\mathrm{D}_m}$, since the Jacobian of the polynomials $N_i \circ F_{\mathrm{D}_m, w}$, $i = 2, 4, \cdots, 2n-2, n$, is up to a constant factor given by $\prod_{i<j}(x_i^2 - x_j^2)$. If $m (\ge 4)$ is even, this cannot be true since there have to be two basic invariants of degree $m/2$.

13.7.5. Theorem. *Suppose $G \subseteq \mathrm{O}(V)$ is a finite irreducible reflection group and $G \ne \mathrm{D}_m$, $m \ge 4$. Let $v \in V \backslash \{0\}$ such that the cardinality of G_v is maximal. Let*

$$P(t)(x) = x^n - a_1(t) x^{n-1} + a_2(t) x^{n-2} - \cdots + (-1)^n a_n(t) \qquad (t \in \mathbb{R})$$

be a smooth curve of hyperbolic polynomials of degree $n = |G.v|$ *lying in* $P_{G,v}(V)$ *for all* $t \in \mathbb{R}$*. Then there exists a global twice differentiable parameterization of the roots of* $P(t)$ *on* $\mathbb{R}$ *which is smooth on* $\mathbb{R}\backslash F_{\mathrm{orb}}$*.*

PROOF. The theorem is a corollary of theorem 13.7.4 and theorem 13.4.2. □

REMARK. It is easy to see that, under the assumption that the cardinality of G_v is maximal, the orbit type stratification and the ambient stratification of $U_{G,v}$ coincide only for $G = \mathrm{A}_m, \mathrm{B}_m, \mathrm{I}_2^m$. In general, if $|G_v|$ is not maximal, the orbit type stratification of $U_{G,v}$ will be strictly coarser than its ambient stratification.

13.7.6. The integer k, associated to orthogonal representations of finite groups G in 13.5.2, has been determined for finite irreducible reflection groups G in figure 12.2.

In the situation of theorem 13.7.5 the strategy discussed in section 13.5 will lead to no improvement, since $k = n$, by definition. But, if we choose $v \in V\backslash\{0\}$ such that $|G_v|$ is not maximal, then $k < n$ and the methods of section 13.5 will yield refinements.

In many cases the following theorem provides an improvement of 13.5.1.

THEOREM. *Suppose* $G \subseteq \mathrm{O}(V)$ *is a finite irreducible reflection group. Choose some* $v \in V\backslash\{0\}$*. Put* $n = |G.v|$ *and let* k *be as in figure 12.2. Suppose that the restrictions* $E_i|_{U_{G,v}}$*,* $1 \le i \le n$*, generate* $\mathbb{R}[U_{G,v}]^{W_{G,v}}$*. Let*

$$P(t)(x) = x^n - a_1(t)x^{n-1} + a_2(t)x^{n-2} - \cdots + (-1)^n a_n(t) \qquad (t \in \mathbb{R})$$

be a curve of hyperbolic polynomials lying in $P_{G,v}(V)$ *for all* $t \in \mathbb{R}$*. Then:*

(1) *If all* a_i *are of class* C^k*, then there exists a differentiable parameterization of the roots of* $P(t)$ *with locally bounded derivative.*
(2) *If all* a_i *are of class* C^{k+d}*, then any differentiable parameterization of the roots of* $P(t)$ *is actually* C^1*.*
(3) *If all* a_i *are of class* C^{k+2d}*, then there exists a twice differentiable parameterization of the roots of* $P(t)$*.* □

13.7.7. EXAMPLE. Consider the Coxeter group B_3 and choose $v = e_1 + e_2 + e_3$. We find

$$\begin{aligned} F_{\mathrm{B}_3,v}(x) = (&x_1 + x_2 + x_3, -x_1 + x_2 + x_3, x_1 - x_2 + x_3, x_1 + x_2 - x_3, \\ &- x_1 - x_2 + x_3, -x_1 + x_2 - x_3, x_1 - x_2 - x_3, -x_1 - x_2 - x_3) \end{aligned}$$

and $U_{\mathrm{B}_3,v} = \{y \in \mathbb{R}^8 : y_i + y_j = 0 \text{ for } i + j = 9, y_1 = y_2 + y_3 + y_4\}$. It is easy to check that $N_{2i} \circ F_{\mathrm{B}_3,v}$, $1 \le i \le 3$, generate $\mathbb{R}[\mathbb{R}^3]^{\mathrm{B}_3}$, by computing their Jacobian. It is readily verified that the set of all reflection hyperplanes of $W_{\mathrm{B}_3,v}$ is given by intersecting the following hyperplanes in $\mathbb{R}^8$ with $U_{\mathrm{B}_3,v}$ (compare with remark 13.7):

$$\{y_1 = y_2, y_1 = y_3, y_1 = y_4, y_1 = y_5, y_1 = y_6, y_1 = y_7, y_2 = y_3, y_2 = y_4, y_3 = y_4\}.$$

Furthermore, the intersections with $U_{\mathrm{B}_3,v}$ of the following hyperplanes in $\mathbb{R}^8$,

$$\{y_1 = y_8, y_2 = y_7, y_3 = y_6, y_4 = y_5\},$$

are not among the set of reflection hyperplanes of $W_{\mathrm{B}_3,v}$. Therefore, the orbit type stratification of $U_{\mathrm{B}_3,v}$ is strictly coarser than its ambient stratification.

We follow the recipe for computing orbit type and ambient stratification of $E(U_{\mathrm{B}_3,v}) = N(U_{\mathrm{B}_3,v})$ given in 13.4.3. We will present only the outcome of the calculations. Using $N_{2i} \circ F_{\mathrm{B}_3,v}$, $1 \le i \le 3$, as basic invariants of $\mathbb{R}[\mathbb{R}^3]^{\mathrm{B}_3}$, we find that the symmetric matrix $\tilde{B} = (\tilde{b}_{ij})$ from (9.10) has entries

$$\begin{aligned}
\tilde{b}_{11} &= 32z_2,\ \tilde{b}_{12} = 64z_4,\ \tilde{b}_{13} = 96z_6,\ \tilde{b}_{22} = -3z_2^3 + 36z_2z_4 + 32z_6,\\
\tilde{b}_{23} &= \frac{1}{8}(5z_2^4 - 108z_2^2z_4 + 192z_4^2 + 544z_2z_6),\\
\tilde{b}_{33} &= \frac{1}{64}(27z_2^5 - 300z_2^3z_4 - 1140z_2z_4^2 + 1140z_2^2z_6 + 7680z_4z_6).
\end{aligned}$$

Put $\tilde{\Delta}_{ij} = \det \begin{pmatrix} \tilde{b}_{ii} & \tilde{b}_{ij} \\ \tilde{b}_{ji} & \tilde{b}_{jj} \end{pmatrix}$ where $i < j$. Then $N(U_{\mathrm{B}_3,v})$ is the subset in $\mathbb{R}^8$ defined by the following relations

$$\begin{gathered}
z_2 \ge 0, \tilde{\Delta}_{12} \ge 0, \det \tilde{B} \ge 0\\
z_1 = z_3 = z_5 = z_7 = 0,\\
384z_8 = 5z_2^4 - 72z_2^2z_4 + 48z_4^2 + 256z_2z_6.
\end{gathered}$$

The 3-dimensional principal orbit type stratum is given by

$$R^{(3)} = N(U_{\mathrm{B}_3,v}) \cap \{z_2 > 0, \tilde{\Delta}_{12} > 0, \det \tilde{B} > 0\}.$$

Put

$$\begin{aligned}
\tilde{f}_1 &= 53z_2^6 - 840z_2^4z_4 + 1680z_2^2z_4^2 + 6144z_4^3 + 2752z_2^3z_6 - 16128z_2z_4z_6 + 9216z_6^2,\\
\tilde{f}_2 &= z_2^3 - 12z_2z_4 + 32z_6.
\end{aligned}$$

There are three 2-dimensional orbit type strata

$$\begin{aligned}
R_1^{(2)} &= N(U_{\mathrm{B}_3,v}) \cap \{z_2 > 0, \tilde{\Delta}_{12} > 0, \tilde{f}_1 = 0\}\\
R_2^{(2)} &= N(U_{\mathrm{B}_3,v}) \cap \{z_2 > 0, \tilde{\Delta}_{12} = 0, \tilde{\Delta}_{23} > 0, \tilde{f}_1 = 0\}\\
R_3^{(2)} &= N(U_{\mathrm{B}_3,v}) \cap \{z_2 > 0, \tilde{\Delta}_{13} > 0, \tilde{f}_2 = 0\},
\end{aligned}$$

the three 1-dimensional orbit type strata $R_1^{(1)}$, $R_2^{(1)}$, $R_3^{(1)}$ are the connected components of

$$N(U_{\mathrm{B}_3,v}) \cap \{z_2 > 0, \tilde{\Delta}_{12} = \tilde{\Delta}_{13} = \tilde{\Delta}_{23} = 0\},$$

and $R^{(0)} = \{0\}$ is the only 0-dimensional stratum.

The ambient stratification of $N(U_{\mathrm{B}_3,v})$ is obtained by cutting with the surface $\{z_2^2 - 4z_4 = 0\}$. There are two 3-dimensional ambient strata

$$S_1^{(3)} = R^{(3)} \cap \{z_2^2 - 4z_4 > 0\} \quad \text{and} \quad S_2^{(3)} = R^{(3)} \cap \{z_2^2 - 4z_4 < 0\},$$

five 2-dimensional ambient strata

$$\begin{aligned}
S_1^{(2)} &= R^{(3)} \cap \{z_2^2 - 4z_4 = 0\},\ S_2^{(2)} = R_1^{(2)} \cap \{z_2^2 - 4z_4 > 0\},\\
S_3^{(2)} &= R_1^{(2)} \cap \{z_2^2 - 4z_4 < 0\},\ S_4^{(2)} = R_2^{(2)},\ S_5^{(2)} = R_3^{(2)},
\end{aligned}$$

four 1-dimensional ambient strata $S_1^{(1)} = R_1^{(1)}$, $S_2^{(1)} = R_2^{(1)}$, $S_3^{(1)} = R_3^{(1)}$, $S_4^{(1)} = R_1^{(2)} \cap \{z_2^2 - 4z_4 = 0\}$, and $S^{(0)} = R^{(0)} = \{0\}$ is the only 0-dimensional ambient stratum. See figure 13.1.

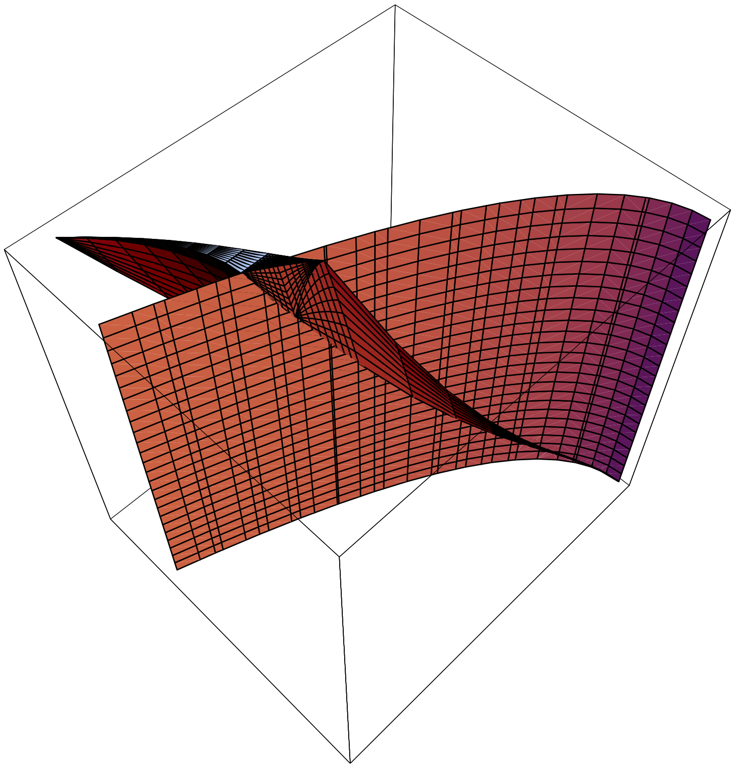

FIGURE 13.1. The projection of $N(U_{\mathrm{B}_3,v})$ to the $\{z_2, z_4, z_6\}$-subspace and intersection with the surface $\{z_2^2 - 4z_4 = 0\}$.

Let f, g, h be functions defined in some neighborhood of $0 \in \mathbb{R}$. Suppose that f and g are infinitely flat at 0 and $h(0) = 0$. For t near 0 consider the curve of polynomials $P(t)(x) = x^8 + \sum_{j=1}^{8} (-1)^j a_j(t) x^{8-j}$ where

$$a_1 = a_3 = a_5 = a_7 = 0,$$
$$a_2 = -56 + f,\ a_4 = 784 + g,\ a_6 = -2304 + h,$$
$$1024 a_8 = 16 a_2^4 - 128 a_2^2 a_4 + 256 a_4^2.$$

Then, for t near 0, $P(t)$ is a curve in $N(U_{\mathrm{B}_3,v})$ with $P(0) \in S_1^{(2)}$. At $t = 0$ it is normally flat with respect to the ambient stratification but normally nonflat with respect to the orbit type stratification.

If f, g and h are smooth, then $P(t)$ is smoothly solvable near $t = 0$, by theorem 13.4.2. Note that in this example we have $d = k = 6 < 8 = n$ and thus theorem 13.7.6 provides an actual improvement, too.

13.7.8. The following example shows that $W(U)$ must not necessarily be a finite reflection group, even though the $E_i|_U$ generate $\mathbb{R}[U]^{W(U)}$.

EXAMPLE. Let U be the subspace of $\mathbb{R}^6$ defined by the following equations

$$x_1 + x_2 + x_3 = 0, \quad x_4 + x_5 + x_6 = 0.$$

The subgroup $N(U)$ of S_6 is generated by all permutations of x_1, x_2, x_3, all permutations of x_4, x_5, x_6, and the simultaneous transpositions of x_1 and x_4, x_2 and x_5, x_3 and x_6. The subgroup $Z(U)$ is trivial. Thus $W(U)$ is isomorphic to the semidirect product of $\mathrm{S}_3 \times \mathrm{S}_3$ and S_2.

One can get the subspace U above as follows. Consider the point $v = (x, x, x, y, y, y) \in \mathbb{R}^6$, where $x, y \neq 0$ and $x \neq y$. The isotropy group $H = (\mathrm{S}_6)_v$ of v is evidently isomorphic to $\mathrm{S}_3 \times \mathrm{S}_3$. Then $U = ((\mathbb{R}^6)^H)^\perp$. The group H is the normal subgroup of $W(U)$ generated by reflections.

First consider the action of H on U. It is clear that the algebra $\mathbb{R}[U]^H$ is a polynomial algebra generated by the basic generators

$$y_1 = x_1^2 + x_2^2 + x_1x_2,\ z_1 = x_1x_2(x_1 + x_2),$$
$$y_2 = x_4^2 + x_5^2 + x_4x_5,\ z_2 = x_4x_5(x_4 + x_5).$$

Consider the space $\mathbb{R}^4$ with the coordinates y_1, z_1, y_2, z_2 and the action of the group S_2 on it induced by the action of $\mathrm{S}_2 = W(U)/(\mathrm{S}_3 \times \mathrm{S}_3)$ on the above basic generators. It is easy to check that this action coincides with the diagonal action of S_2 on $(\mathbb{R}^2)^2$ for the standard action of S_2 on $\mathbb{R}^2$. Since the algebra of S_2-invariant polynomials on $(\mathbb{R}^2)^2$ is generated by the polarizations of basic invariants for the standard action of S_2 ob $\mathbb{R}^2$ we get the following system of generators of $\mathbb{R}[U]^{W(U)}$:

$$f_1 = y_1 + y_2,\ f_2 = z_1 + z_2,\ f_3 = y_1^2 + y_2^2,\ f_4 = y_1z_1 + y_2z_2,\ f_5 = z_1^2 + z_2^2.$$

Simple calculations for the restrictions of the Newton polynomials N_i on $\mathbb{R}^6$ to U gives the following result:

$$N_1|_U = 0,\ N_2|_U = 2f_1,\ N_3|_U = -3f_2,$$
$$N_4|_U = 2f_3,\ N_5|_U = -5f_4,\ N_6|_U = 3f_5 + 3f_1f_3 - f_1^3.$$

This proves that the morphism $\mathbb{R}[\mathbb{R}^6]^{\mathrm{S}_6} \to \mathbb{R}[U]^{W(U)}$ defined by restriction is surjective.

Appendix

13.8. Basic invariants of finite reflection groups and finite rotation groups

We shall list in this section minimal sets of basic invariants of the irreducible Coxeter groups and of the finite rotation groups in two and three dimensions. They are not unique. However, their degrees are unique, e.g. 11.2.1, or [**GB85**], [**Hum90**].

13.8.1. Basic invariants for irreducible Coxeter groups.

Coxeter group	Degrees of basic invariants
A_n, $n \geq 1$	$2, 3, \ldots, n+1$
B_n, $n \geq 2$	$2, 4, \ldots, 2n$
D_n, $n \geq 4$	$2, 4, \ldots, 2n-2, n$
I_2^n, $n \geq 5$	$2, n$
G_2	$2, 6$
H_3	$2, 6, 10$
H_4	$2, 12, 20, 30$
F_4	$2, 6, 8, 12$
E_6	$2, 5, 6, 8, 9, 12$
E_7	$2, 6, 8, 10, 12, 14, 18$
E_8	$2, 8, 12, 14, 18, 20, 24, 30$

FIGURE 13.2. Classification of irreducible Coxeter groups and corresponding degrees of basic invariants

The basic invariants for the irreducible Coxeter groups listed below are taken from [**PSW78**].

A_n: We may view the symmetric group S_{n+1} as a group of linear transformations of $\mathbb{R}^{n+1}$, if we agree that each element of S_{n+1} permutes the basis vectors $e_1, e_2, \ldots, e_{n+1}$. Let V denote the hyperplane in $\mathbb{R}^{n+1}$ given by

$$x_1 + x_2 + \cdots + x_{n+1} = 0.$$

Then A_n is the group of restrictions to V of the transformations in S_{n+1}. As generators of the algebra $\mathbb{R}[V]^{\mathrm{A}_n}$ of A_n-invariant polynomials may be taken

$$\sigma_i(x_1, \ldots, x_{n+1}) = \sum_{j=1}^{n} x_j^i \qquad (2 \le i \le n+1)$$
$$(x_1 + x_2 + \cdots + x_{n+1} = 0).$$

B_n: For B_n we may consider the group of 'signed permutations' in $\mathbb{R}^n$ whose elements permute the basis vectors $e_1, e_2, \ldots, e_n$, and then replace some of them by their negatives. More precisely, let K_n denote the group generated by the n reflections along the basis vectors $e_1, \ldots, e_n$ and S_n the symmetric group as above. Then B_n is the semidirect product of K_n by S_n:

$$\mathrm{B}_n = \mathrm{K}_n \rtimes \mathrm{S}_n .$$

As generators of the algebra $\mathbb{R}[\mathbb{R}^n]^{\mathrm{B}_n}$ of B_n-invariant polynomials may be taken

$$\sigma_i(x_1, \ldots, x_n) = \sum_{j=1}^{n} x_j^{2i} \qquad (1 \le i \le n).$$

D_n: For D_n we may consider the group of all permutations of the basis vectors $e_1, \ldots, e_n$ in $\mathbb{R}^n$ followed by even numbers of sign changes. We replace K_n in B_n by its subgroup L_n consisting of products of even-numbered reflections along the basis vectors $e_1, \ldots, e_n$. Then D_n is the semidirect product of L_n by S_n:

$$\mathrm{D}_n = \mathrm{L}_n \rtimes \mathrm{S}_n .$$

As generators of the algebra $\mathbb{R}[\mathbb{R}^n]^{\mathrm{D}_n}$ of D_n-invariant polynomials may be taken

$$\sigma_i(x_1, \ldots, x_n) = \sum_{j=1}^{n} x_j^{2i} \qquad (1 \le i \le n-1)$$
$$\sigma_n(x_1, \ldots, x_n) = x_1 x_2 \cdots x_n .$$

I_2^n: The dihedral group I_2^n (for $n \ge 3$) consists of all orthogonal transformations which preserve a regular n-sided polygon centered at 0 in the Euclidean plane $\mathbb{R}^2$. As generators of the algebra $\mathbb{R}[\mathbb{R}^2]^{\mathrm{I}_2^n}$ of I_2^n-invariant polynomials may be taken

$$\sigma_1(x, y) = z\bar{z} = x^2 + y^2$$
$$\sigma_2(x, y) = \mathrm{Re}(z^n) = \sum_{j=0}^{\left[\frac{n}{2}\right]} (-1)^j \binom{n}{2j} x^{n-2j} y^{2j},$$

where $z = x + iy$, and $[r]$ is the largest integer less or equal r.

G_2: The group G_2 is isomorphic to I_2^6. As basic invariant we may choose

$$\sigma_1(x, y) = x^2 + y^2$$
$$\sigma_2(x, y) = x^6 - 15x^4y^2 + 15x^2y^4 - y^6.$$

H_3: The group H_3 can be taken as the one generated by reflections in the planes $x_2 = 0$, $x_3 = 0$, and $\tau^{-1}x_1 - x_2 - \tau x_3 = 0$ in $\mathbb{R}^3$, where $\tau = \frac{1}{2}(\sqrt{5} + 1)$ is the golden ratio. It is the symmetry group of the icosahedron and dodecahedron. Put

$$I_{2k}(x_1, x_2, x_3) = \sum_{j=0}^{k} \binom{2k}{2j} \tau^{2j} \left(x_1^{2j} x_2^{2k-2j} + x_2^{2j} x_3^{2k-2j} + x_3^{2j} x_1^{2k-2j} \right).$$

Then $\sigma_1 = I_2$, $\sigma_2 = I_6$, and $\sigma_3 = I_{10}$ generate $\mathbb{R}[\mathbb{R}^3]^{\mathrm{H}_3}$.

H_4: The group H_4 can be defined as the one generated by reflections in the hyperplanes $x_3 = 0$, $x_4 = 0$, $\tau^{-1}x_2 - \tau x_3 - x_4 = 0$, and $\tau^{-1}x_1 - \tau x_2 - x_4 = 0$ in $\mathbb{R}^4$, where $\tau = \frac{1}{2}(\sqrt{5}+1)$ is the golden ratio. It is the symmetry group of the 600-cell and 120-cell. The set of the 120 linear forms

$$\begin{aligned} &\pm 2x_i, \qquad 1 \le i \le 4,\\ &\pm x_1 \pm x_2 \pm x_3 \pm x_4,\\ &\pm \tau x_1 \pm \tau^{-1}x_2 \pm x_3,\\ &\pm \tau x_1 \pm \tau^{-1}x_3 \pm x_4,\\ &\pm \tau x_1 \pm \tau^{-1}x_4 \pm x_2,\\ &\pm \tau x_2 \pm \tau^{-1}x_4 \pm x_3, \end{aligned}$$

where in the last 4 rows $(\tau, \tau^{-1}, 1)$ is cyclically permuted, is invariant under H_4. Let I_j be the sum of the j-th powers of these linear form. Then $\sigma_1 = I_2$, $\sigma_2 = I_{12}$, $\sigma_3 = I_{20}$, and $\sigma_4 = I_{30}$ generate $\mathbb{R}[\mathbb{R}^4]^{\mathrm{H}_4}$.

F_4: The group F_4 can be taken as the one generated by reflections in the planes $x_2 - x_3 = 0$, $x_3 - x_4 = 0$, $x_4 = 0$, and $x_1 - x_2 - x_3 - x_4 = 0$ in $\mathbb{R}^4$. It is the symmetry group of the 24-cell. Put

$$I_{2k}(x_1, x_2, x_3, x_4) = \sum_{1\le i<j\le 4} \left((x_i + x_j)^{2k} + (x_i - x_j)^{2k}\right).$$

We have $I_{2k} = (8 - 2^{2k-1})s_{2k} + \sum_{j=1}^{k-1} \binom{2k}{2j} s_{2j}s_{2k-2j}$, where s_i is the i-th Newton polynomial. Then $\sigma_1 = I_2$, $\sigma_2 = I_6$, $\sigma_3 = I_8$, and $\sigma_4 = I_{12}$ generate $\mathbb{R}[\mathbb{R}^4]^{\mathrm{F}_4}$.

E_6: Let us consider the group generated by reflections in the six hyperplanes $x_1 - x_2 = 0$, $x_2 - x_3 = 0$, $x_3 - x_4 = 0$, $x_4 - x_5 = 0$, $x_5 - x_6 = 0$, and $x_1 + x_2 + x_3 - x_4 - x_5 - x_6 + x_7 - x_8 = 0$ in $\mathbb{R}^8$. Restricting this group to the 6-dimensional subspace $\{x_1 + \cdots + x_6 = 0, x_7 + x_8 = 0\}$ yields E_6. Define $\frac{1}{2}(x_7 - x_8) = x_7 =: y$ and

$$I_k(x_1, \dots, x_6, y) = \sum_{1\le i<j\le 6} (x_i + x_j)^k + \sum_{i=1}^{6} \left((-x_i + y)^k + (-x_i - y)^k\right).$$

If s_i is the i-th Newton polynomial in $x_1, \dots, x_6$, then

$$I_k = (6 - 2^{k-1})s_k + \frac{1}{2}\sum_{j=2}^{k-2} \binom{k}{j} s_j s_{k-j} + (-1)^k 2 \sum_{j=0}^{\left[\frac{k}{2}\right]} \binom{k}{2j} y^{2j} s_{k-2j}.$$

Then $\sigma_1 = I_2$, $\sigma_2 = I_5$, $\sigma_3 = I_6$, $\sigma_4 = I_8$, $\sigma_5 = I_9$, and $\sigma_6 = I_{12}$ generate the algebra of E_6-invariant polynomials.

E_7: Let us consider the group generated by reflections in the seven hyperplanes $x_1 - x_2 = 0$, $x_2 - x_3 = 0$, $x_3 - x_4 = 0$, $x_4 - x_5 = 0$, $x_5 - x_6 = 0$, $x_6 - x_7 = 0$ and $x_1 + x_2 + x_3 + x_4 - x_5 - x_6 - x_7 - x_8 = 0$ in $\mathbb{R}^8$. Restricting this group to the hyperplane $x_1 + \cdots + x_8 = 0$ yields E_7. Define

$$I_{2k}(x_1, \dots, x_8) = 2 \sum_{1\le i<j\le 8} (x_i + x_j)^{2k}.$$

If s_i is the i-th Newton polynomial in $x_1, \dots, x_8$, then

$$I_{2k} = (16 - 2^{2k})s_{2k} + \sum_{j=2}^{2k-2} \binom{2k}{j} s_j s_{2k-j}.$$

Then $\sigma_1 = I_2$, $\sigma_2 = I_6$, $\sigma_3 = I_8$, $\sigma_4 = I_{10}$, $\sigma_5 = I_{12}$, $\sigma_6 = I_{14}$, and $\sigma_7 = I_{18}$ generate the algebra of E_7-invariant polynomials.

E_8: Let us consider the group generated by reflections in the eight hyperplanes $x_1 - x_2 = 0$, $x_2 - x_3 = 0$, $x_3 - x_4 = 0$, $x_4 - x_5 = 0$, $x_5 - x_6 = 0$, $x_6 - x_7 = 0$, $x_7 - x_8 = 0$, and $2x_1 + 2x_2 + 2x_3 - x_4 - x_5 - x_6 - x_7 - x_8 - x_9 = 0$ in $\mathbb{R}^9$. Restricting this group to the hyperplane $x_1 + \cdots + x_9 = 0$ yields E_8. Define

$$I_{2k}(x_1, \ldots, x_9) = \sum_{1 \leq i < j \leq 9} (x_i - x_j)^{2k} + \sum_{1 \leq i < j < l \leq 9} (x_i + x_j + x_l)^{2k}.$$

If s_i is the i-th Newton polynomial in $x_1, \ldots, x_9$, then

$$I_{2k} = 9(3^{2k-3} - 2^{2k-1} + 5)s_{2k} + \frac{1}{2} \sum_{j=2}^{2k-2} (9 + (-1)^k - 2^k) \binom{2k}{j} s_j s_{2k-j}$$
$$+ \frac{1}{6} \sum_{\substack{i,j \geq 2 \\ i+j \leq 2k-2}} \frac{(2k)!}{i!j!(2l - i - j)!} s_i s_j s_{2k-i-j}.$$

Then $\sigma_1 = I_2$, $\sigma_2 = I_8$, $\sigma_3 = I_{12}$, $\sigma_4 = I_{14}$, $\sigma_5 = I_{18}$, $\sigma_6 = I_{20}$, $\sigma_7 = I_{24}$, and $\sigma_8 = I_{30}$ generate the algebra of E_8-invariant polynomials.

13.8.2. Basic invariants for finite rotation groups in dimensions two and three.

Rotation group	Degrees of basic invariants
C_2^n, $n \geq 1$	$2, n, n$
C_3^n, $n \geq 1$	$1, 2, n, n$
I_3^n, $n \geq 2$	$2, 2, n, n+1$
T	$2, 3, 4, 6$
W	$2, 4, 6, 9$
H	$2, 6, 10, 15$

FIGURE 13.3. Classification of the finite rotation groups in two and three dimensions and corresponding degrees of basic invariants.

C_2^n: As generators for $\mathbb{R}[\mathbb{R}^2]^{\mathrm{C}_2^n}$ may be taken

$$\sigma_1(x, y) = z\bar{z} = x^2 + y^2$$
$$\sigma_2(x, y) = \mathrm{Re}(z^n) = \sum_{j=0}^{[\frac{n}{2}]} (-1)^j \binom{n}{2j} x^{n-2j} y^{2j}$$
$$\sigma_3(x, y) = \mathrm{Im}(z^n) = \sum_{j=0}^{[\frac{n+1}{2}]-1} (-1)^j \binom{n}{2j+1} x^{n-2j-1} y^{2j+1},$$

where $z = x + iy$, and $[r]$ is the largest integer less or equal r. There is the following relation $\sigma_1^n = \sigma_2^2 + \sigma_3^2$.

C_3^n: It follows from the construction of the group C_3^n that the following polynomials generate $\mathbb{R}[\mathbb{R}^3]^{C_3^n}$:

$$\sigma_1(x,y,z) = x^2 + y^2$$
$$\sigma_2(x,y,z) = \sum_{j=0}^{\left[\frac{n}{2}\right]} (-1)^j \binom{n}{2j} x^{n-2j} y^{2j}$$
$$\sigma_3(x,y,z) = \sum_{j=0}^{\left[\frac{n+1}{2}\right]-1} (-1)^j \binom{n}{2j+1} x^{n-2j-1} y^{2j+1}$$
$$\sigma_4(x,y,z) = z.$$

We have again $\sigma_1^n = \sigma_2^2 + \sigma_3^2$.

I_3^n: It is easy to verify that the following polynomials generate $\mathbb{R}[\mathbb{R}^3]^{I_3^n}$:

$$\sigma_1(x,y,z) = x^2 + y^2$$
$$\sigma_2(x,y,z) = \sum_{j=0}^{\left[\frac{n}{2}\right]} (-1)^j \binom{n}{2j} x^{n-2j} y^{2j}$$
$$\sigma_3(x,y,z) = z^2$$
$$\sigma_4(x,y,z) = z \sum_{j=0}^{\left[\frac{n+1}{2}\right]-1} (-1)^j \binom{n}{2j+1} x^{n-2j-1} y^{2j+1}.$$

They fulfill the following relation

$$\sigma_1^n \sigma_3 - \sigma_2^2 \sigma_3 - \sigma_4^2 = 0.$$

T: As generators of the algebra $\mathbb{R}[\mathbb{R}^3]^{T}$ of T-invariant polynomials on $\mathbb{R}^3$ may be taken (see e.g. [**CP88**])

$$\sigma_1(x,y,z) = x^2 + y^2 + z^2$$
$$\sigma_2(x,y,z) = x^4 + y^4 + z^4$$
$$\sigma_3(x,y,z) = xyz$$
$$\sigma_4(x,y,z) = (x^2 - y^2)(y^2 - z^2)(z^2 - x^2).$$

These basic invariants satisfy the following relation

$$\sigma_1^6 - 4\sigma_1^4\sigma_2 + 5\sigma_1^2\sigma_2^2 - 2\sigma_2^3 - 20\sigma_1^3\sigma_3^2 + 36\sigma_1\sigma_2\sigma_3^2 + 108\sigma_3^4 + 4\sigma_4^2 = 0.$$

W: A system of generators of $\mathbb{R}[\mathbb{R}^3]^{W}$ is given by (see e.g. [**PSW78**])

$$\sigma_1(x,y,z) = x^2 + y^2 + z^2$$
$$\sigma_2(x,y,z) = x^4 + y^4 + z^4$$
$$\sigma_3(x,y,z) = x^6 + y^6 + z^6$$
$$\sigma_4(x,y,z) = xyz(x^2 - y^2)(y^2 - z^2)(z^2 - x^2).$$

These basic invariants satisfy the following relation

$$\begin{aligned} 0 = {} & \sigma_1^9 - 12\sigma_1^7\sigma_2 + 48\sigma_1^5\sigma_2^2 - 66\sigma_1^3\sigma_2^3 + 9\sigma_1\sigma_2^4 \\ & + 10\sigma_1^6\sigma_3 - 78\sigma_1^4\sigma_2\sigma_3 + 150\sigma_1^2\sigma_2^2\sigma_3 - 6\sigma_2^3\sigma_3 + 34\sigma_1^3\sigma_3^2 \\ & - 126\sigma_1\sigma_2\sigma_3^2 + 36\sigma_3^3 + 36\sigma_4^2. \end{aligned}$$

H: We refer to [**CP88**]. Let $\tau = \frac{1}{2}(\sqrt{5}+1)$ be the golden ratio, let s_i denote the i-th Newton polynomial in x, y, z, and put $a_3 = xyz$ and $b_6 = (x^2-y^2)(y^2-z^2)(z^2-x^2)$. Then a system of generators of $\mathbb{R}[\mathbb{R}^3]^{\mathrm{H}}$ is given by

$$\begin{aligned}
\sigma_1 &= s_2 \\
\sigma_2 &= (4\tau - 2)b_6 + 22a_3^2 + s_2 s_4 \\
\sigma_3 &= 3\sigma_2 s_4 - 8s_2 s_4^2 + 9s_2^3 s_4 - 256a_3^2 s_4 + 128a_3^2 s_2^2 \\
\sigma_4 &= \sigma_2(15a_3 s_2 s_4 + 290a_3^3 - 11a_3 s_2^3) - 225a_3 s_4^3 + 425a_3 s_2^2 s_4^2 \\
&\quad - 80a_3^3 s_2 s_4 - 270a_3 s_2^4 s_4 - 9728a_3^5 + 54a_3^3 s_2^3 + 58a_3 s_2^6.
\end{aligned}$$

These basic invariants satisfy the following relation

$$\begin{aligned}
0 &= 80\sigma_4^2 + 50\sigma_3^3 - 550\sigma_1^2\sigma_2\sigma_3^2 - 66\sigma_1^5\sigma_3^2 + 450\sigma_1\sigma_2^3\sigma_3 \\
&\quad + 360\sigma_1^4\sigma_2^2\sigma_3 + 2458\sigma_1^7\sigma_2\sigma_3 - 740\sigma_1^{10}\sigma_3 - 135\sigma_2^5 - 215\sigma_1^3\sigma_2^4 \\
&\quad - 1200\sigma_1^6\sigma_2^3 - 776\sigma_1^9\sigma_2^2 - 2625\sigma_1^{12}\sigma_2 + 1495\sigma_1^{15}.
\end{aligned}$$

Part 3

Related topics

CHAPTER 14

Perturbation of linear operators

The perturbation problem for hyperbolic polynomials is in a natural way related to the perturbation problem for linear operators. We give some overview which is based on [**AKLM98**] and [**KM03**].

14.1. Choosing eigenvalues and eigenvectors of matrices smoothly

14.1.1. Classical results. Let $A(t) = (A_{ij}(t))$ be a smooth (real analytic, holomorphic) curve of real (complex) $(n \times n)$-matrices or linear operators, depending on a real (complex) parameter t near 0. What can we say about the eigenvalues and eigenvectors of $A(t)$?

Let us recall some classical results.

THEOREM ([**Kat76**, II.1.8]). *Let $\mathbb{C} \ni t \mapsto A(t)$ be a holomorphic curve. Then all eigenvalues, all eigenprojections, and all eigennilpotents are holomorphic with at most algebraic singularities at discrete points.*

THEOREM ([**Rel37**, Satz 1]). *Let $t \mapsto A(t)$ be a real analytic curve of Hermitian complex matrices. Let λ be a k-fold eigenvalue of $A(0)$ with k orthonormal eigenvectors v_i, and suppose that there is no other eigenvalue of A near λ. Then there are k real analytic eigenvalues λ_i through λ, and k orthonormal real analytic eigenvectors through the v_i, for t near 0.*

The condition that $A(t)$ is Hermitian cannot be omitted, as shown by:

EXAMPLE. Consider

$$A(t) = \begin{pmatrix} 2t + t^3 & t \\ -t & 0 \end{pmatrix}$$

with eigenvalues $\lambda_\pm(t) = t + \frac{t^3}{2} \pm t^2\sqrt{1 + \frac{t^2}{4}}$ and eigenvectors $x_\pm(t) = (1 + \frac{t^2}{2} \pm t\sqrt{1 + \frac{t^2}{4}}, -1)$ which do not provide a base at $t = 0$.

THEOREM ([**Kat76**, II.6.8], [**Rel69**]). *Let $A(t)$ be a C^1 curve of symmetric matrices. Then the eigenvalues can be chosen C^1 in t, on the whole parameter interval.*

This result is best possible for the degree of continuous differentiability, as proves the following example.

EXAMPLE. We consider the symmetric matrix

$$A(t) = \begin{pmatrix} a(t) & b(t) \\ b(t) & -a(t) \end{pmatrix}$$

with characteristic polynomial $\lambda^2 - (a(t)^2 + b(t)^2)$. We shall specify the entries a and b as smooth functions in such a way that $a(t)^2 + b(t)^2$ does not admit a C^2 square root.

Assume that $a(t)^2 + b(t)^2 = c(t)^2$ for a C^2 function c. Then:

$$\begin{aligned}
c^2 &= a^2 + b^2 \\
cc' &= aa' + bb' \\
(c')^2 + cc'' &= (a')^2 + aa'' + (b')^2 + bb'' \\
c'' &= \frac{(ab' - ba')^2 + a^3a'' + b^3b'' + ab^2a'' + a^2bb''}{c^3}
\end{aligned}$$

Since $c^2 = a^2 + b^2$ we have

$$\left|\frac{a^3}{c^3}\right| \le 1, \quad \left|\frac{b^3}{c^3}\right| \le 1, \quad \left|\frac{ab^2}{c^3}\right| \le \frac{1}{\sqrt{3}}, \quad \left|\frac{a^2b}{c^3}\right| \le \frac{1}{\sqrt{3}}.$$

So for C^2 functions a, b, and continuous c all these terms are bounded. We will now construct smooth functions a and b such that

$$\left(\frac{(ab' - ba')^2}{c^3}\right)^2 = \frac{(ab' - ba')^4}{(a^2 + b^2)^3}$$

is unbounded near $t = 0$. This contradicts that c is C^2. For this we choose a and b similar to the function f in (4.5) in 4.1.4 with the same t_n and h_n:

$$\begin{aligned}
a(t) &:= \sum_{n=1}^{\infty} h_n(t - t_n)\left(\frac{2n}{2^n}(t - t_n) + \frac{1}{4^n}\right), \\
b(t) &:= \sum_{n=1}^{\infty} h_n(t - t_n)\left(\frac{2n}{2^n}(t - t_n)\right).
\end{aligned}$$

Then $a(t_n) = \frac{1}{4^n}$, $b(t_n) = 0$, $|c(t_n)| = \frac{1}{4^n}$, and $b'(t_n) = \frac{2n}{2^n}$.

14.1.2. Theorem ([**AKLM98**]). *Let $A(t) = (A_{ij}(t))$ be a smooth curve of complex Hermitian $(n \times n)$-matrices, depending on a real parameter $t \in \mathbb{R}$, acting on a Hermitian space $V = \mathbb{C}^n$, such that no two of the continuous eigenvalues meets of infinite order at any $t \in \mathbb{R}$ if they are not equal for all t. Then all eigenvalues and all eigenvectors can be chosen smoothly in t, on the whole parameter domain $\mathbb{R}$.*

Proof. We know from theorem 4.3.1 that the characteristic polynomial

$$\chi_{A(t)}(\lambda) = \det(A(t) - \lambda\mathbb{I}) = \sum_{i=0}^{n} \operatorname{trace}(\Lambda^i A(t))\lambda^{n-i}$$

is smoothly solvable, with smooth roots $\lambda_1(t), \dots, \lambda_n(t)$, on the whole parameter interval. Hence we get smooth parameterizations $\lambda_1(t), \dots, \lambda_n(t)$ of the eigenvalues of $A(t)$.

Next we shall show that one can choose the eigenvectors $x_i(t)$ of $A(t)$ smoothly, locally in t. It suffices to show this for t near 0. We proceed by the following algorithm.

Case 1. Suppose that not all eigenvalues of $A(0)$ coincide. Then we can reorder them in such a way that for $i_0 = 0 < 1 \le i_1 < i_2 < \dots < i_k < n = i_{k+1}$ we have

$$\lambda_1(0) = \dots = \lambda_{i_1}(0) < \lambda_{i_1+1}(0) = \dots = \lambda_{i_2}(0) < \dots < \lambda_{i_k+1}(0) = \dots = \lambda_n(0).$$

For t near 0 we still have

$$\lambda_1(t), \dots, \lambda_{i_1}(t) < \lambda_{i_1+1}(t), \dots, \lambda_{i_2}(t) < \dots < \lambda_{i_k+1}(t), \dots, \lambda_n(t).$$

For $1 \le j \le k+1$ consider the subspaces

$$V_t^{(j)} = \bigoplus_{i=i_{j-1}+1}^{i_j} \{v \in V : A(t) = \lambda_i(t)v\}.$$

We claim that each $V_t^{(j)}$ runs through a smooth vector subbundle of the trivial bundle $(-\epsilon,\epsilon)\times V\to(-\epsilon,\epsilon)$, which admits a smooth framing $e_{i_{j-1}+1}(t),\dots,e_{i_j}(t)$, and we have $V=\bigoplus_{j=1}^{k+1}V_t^{(j)}$ for each t.

In order to prove this statement note that

$$V_t^{(j)}=\ker\Big((A(t)-\lambda_{i_{j-1}+1}(t))\circ\dots\circ(A(t)-\lambda_{i_j}(t))\Big),$$

so $V_t^{(j)}$ is the kernel of a smooth vector bundle homomorphism $B(t)$ of constant rank (even of constant dimension of the kernel), and thus is a smooth vector subbundle. This together with a smooth frame field can be shown as follows: Choose a basis of V such that $A(0)$ is diagonal. Then by the elimination procedure one can construct a basis for the kernel of $B(0)$. For t near 0, the elimination procedure (with the same choices) gives then a basis of the kernel of $B(t)$. The elements in that basis are then smooth in t, for t near 0.

From the last result it follows that it suffices to find smooth eigenvectors in each subbundle $V^{(j)}$ separately, expanded in the smooth frame field. But in this frame field the vector subbundle looks again like a constant vector space. So feed each of these parts (A restricted to $V^{(j)}$, as matrix with respect to the frame field) into case 2 below.

Case 2. Assume that all eigenvalues of $A(0)$ are equal. That means that $A(t):V\to V$ is Hermitian with all eigenvalues at $t=0$ equal to $-\frac{1}{n}\operatorname{trace}(\Lambda^1A(0))$. Eigenvectors of $A(t)$ are also eigenvectors of $A(t)+\frac{1}{n}\operatorname{trace}(\Lambda^1A(t))\mathbb{I}$, thus we may replace $A(t)$ by $A(t)+\frac{1}{n}\operatorname{trace}(\Lambda^1A(t))\mathbb{I}$ and assume that the first coefficient of the characteristic polynomial vanishes, or otherwise put, that $\lambda_i(0)=0$ for all i and so $A(0)=0$.

If $A(t)=0$ for all t we choose the eigenvectors constant.

Otherwise, we write $A_{ij}(t)=tA_{ij}^{(1)}(t)$. It follows that the characteristic polynomial of the Hermitian matrix $A^{(1)}(t)$ is

$$\chi_{A^{(1)}(t)}(\lambda)=\sum_{i=0}^{n}t^{-i}\operatorname{trace}(\Lambda^iA(t))\lambda^{n-i}.$$

Compare with section 4.2. The eigenvalues of $A^{(1)}(t)$ are the roots of $\chi_{A^{(1)}(t)}$, which can be chosen smoothly, by theorem 4.3.1. The eigenvectors of $A^{(1)}$ are also eigenvectors of A. If the eigenvalues are still all equal, we apply the same procedure again, until there are distinct eigenvalues. This situation occurs after finitely many steps by the assumptions of the theorem. Then we apply case 1.

This algorithm shows that one may choose the eigenvectors $x_i(t)$ of $A(t)$ in a smooth way, locally in t. It remains to extend this to the whole parameter interval.

If some eigenvalues coincide locally, then on the whole of $\mathbb{R}$, by the assumption. The corresponding eigenspaces then form a smooth vector bundle over $\mathbb{R}$, by case 1, since those eigenvalues which meet in isolated points are different after application of case 2. So we get $V=\bigoplus W_t^{(j)}$ where each $W_t^{(j)}$ is a smooth vector subbundle of $\mathbb{R}\times V$, whose dimension is the generic multiplicity of the corresponding smooth eigenvalue function. It suffices to find global orthonormal smooth frames for each of these vector bundles. But their existence follows from the fact that each such vector bundle is smoothly trivial, by using parallel transport with respect to a smooth Hermitian connection. $\square$

This result cannot be improved:

EXAMPLE ([**Rel37**]). Let $x_+(t) = (\cos\frac{1}{t}, \sin\frac{1}{t})$, $x_-(t) = (-\sin\frac{1}{t}, \cos\frac{1}{t})$, $\lambda_\pm(t) = \pm e^{-\frac{1}{t^2}}$, and

$$\begin{aligned} A(t) &= (x_+(t), x_-(t)) \begin{pmatrix} \lambda_+(t) & 0 \\ 0 & \lambda_-(t) \end{pmatrix} (x_+(t), x_-(t))^{-1} \\ &= e^{-\frac{1}{t^2}} \begin{pmatrix} \cos\frac{2}{t} & \sin\frac{2}{t} \\ \sin\frac{2}{t} & -\cos\frac{2}{t} \end{pmatrix}. \end{aligned}$$

Here $t \mapsto A(t)$ and $t \mapsto \lambda_\pm(t)$ are smooth, whereas the eigenvectors $(\cot\frac{1}{t}, 1)$, $(-\tan\frac{1}{t}, 1)$ cannot be chosen continuously.

14.1.3. REMARK. It is proved in [**KP08**] that each multiparameter real analytic family of symmetric matrices can be diagonalized real analytically after suitable blowing-ups of the space of parameters.

14.2. Perturbation of unbounded compact operators

14.2.1. Curves of unbounded self-adjoint operators. In this section that $A(t)$ is a real analytic, smooth, or $C^{k,\alpha}$ curve of unbounded self-adjoint operators means the following: There is a dense subspace V of the Hilbert space H such that V is the domain of definition of each $A(t)$, and such such that $A(t)^* = A(t)$. Moreover, we require that $t \mapsto \langle A(t)u \mid v\rangle$ is real analytic, smooth, or $C^{k,\alpha}$ for each $u \in V$ and $v \in H$. This implies that $t \mapsto A(t)u$ is of the same class $\mathbb{R} \to H$ for each $u \in V$, by [**KM97**, 2.3]. This is true because $C^{k,\alpha}$ can be described by boundedness conditions only; and for these the uniform boundedness principle is valid. A function f is called $C^{k,\alpha}$ if it is k times differentiable and for the k-th derivative the expression $\frac{f^{(k)}(t)-f^{(k)}(s)}{|t-s|^\alpha}$ is locally bounded in $t \neq s$.

We say that a sequence of continuous, real analytic, smooth, or twice differentiable functions λ_i *parameterize the eigenvalues*, if for each $z \in \mathbb{R}$ the cardinality $|\{i : \lambda_i(t) = z\}|$ equals the multiplicity of z as eigenvalue of $A(t)$.

The following result is well-known.

THEOREM ([**Kat76**, VII.3.9], [**Rel42**]). *Let $A(t)$ be a real analytic curve of unbounded self-adjoint operators in a Hilbert space with common domain of definition and with compact resolvent. Then the eigenvalues and the eigenvectors can be chosen real analytically in t, on the whole parameter domain.*

14.2.2. THEOREM ([**AKLM98**] and [**KM03**]). *Let $t \mapsto A(t)$ be a curve of unbounded self-adjoint operators in a Hilbert space with common domain of definition and with compact resolvent.*

(1) *If $A(t)$ is smooth in $t \in \mathbb{R}$ and if no two of the continuously parameterized eigenvalues meet of infinite order at any $t \in \mathbb{R}$, if they are not equal, then the eigenvalues and the eigenvectors can be chosen smoothly in t, on the whole parameter domain.*
(2) *If $A(t)$ is smooth, then the eigenvalues of $A(t)$ may be parameterized twice differentiably in t.*
(3) *If $A(t)$ is $C^{1,\alpha}$ for some $\alpha > 0$, then the eigenvalues of $A(t)$ may be parameterized in a C^1 way in t*

We are going to prove that theorem in this section.

REMARKS. (1): There are nice applications of theorem 14.2.2:

Let M be a compact manifold and let $t \mapsto g_t$ be a smooth curve of smooth Riemannian metrics on M. Then we get the corresponding smooth curve $t \mapsto \Delta(g_t)$ of Laplace-Beltrami operators on $L^2(M)$. By theorem 14.2.2(2), its eigenvalues can be arranged twice differentiably.

Let Ω be a bounded region in $\mathbb{R}^n$ with smooth boundary, and let $H(t) = -\Delta + V(t)$ be a $C^{1,\alpha}$ curve of Schrödinger operators with varying potential and Dirichlet boundary conditions. Then the eigenvalues can be arranged C^1, by theorem 14.2.2(3).

(2): The following example shows that the conclusion in theorem 14.2.2(2) is best possible: Let $S(2)$ denote the vector space of all symmetric real (2×2)-matrices. We use the general curve lemma [**KM97**, 12.2]: *There exists a converging sequence of reals t_n with the following property: Let $A_n \in C^\infty(\mathbb{R}, S(2))$ be a sequence of functions which converges fast to 0, i.e., for each $k \in \mathbb{N}$ the sequence $n^k A_n$ is bounded in $C^\infty(\mathbb{R}, S(2))$. Then there exists a smooth curve $A \in C^\infty(\mathbb{R}, S(2))$ such that $A(t_n + s) = A_n(s)$ for $|s| \le \frac{1}{n^2}$, for all n.*

Let us use this for

$$A_n(t) := \begin{pmatrix} \frac{1}{2^{n^2}} & \frac{t}{2^n} \\ \frac{t}{2^n} & -\frac{1}{2^{n^2}} \end{pmatrix} = \frac{1}{2^{n^2}} \begin{pmatrix} 1 & \frac{t}{s_n} \\ \frac{t}{s_n} & -1 \end{pmatrix}, \quad \text{where } s_n := 2^{n-n^2} \le \frac{1}{n^2}.$$

The eigenvalues of $A_n(t)$ and their derivatives are

$$\lambda_n(t) = \pm \frac{1}{2^{n^2}} \sqrt{1 + (\frac{t}{s_n})^2}, \quad \lambda_n'(t) = \pm \frac{2^{n^2-2n} t}{\sqrt{1 + (\frac{t}{s_n})^2}}.$$

Then, for the eigenvalues λ of A,

$$\begin{aligned} \frac{\lambda'(t_n + s_n) - \lambda'(t_n)}{s_n^\alpha} &= \frac{\lambda_n'(s_n) - \lambda_n'(0)}{s_n^\alpha} = \pm \frac{2^{n^2-2n} s_n}{s_n^\alpha \sqrt{2}} \\ &= \pm \frac{2^{n(\alpha(n-1)-1)}}{\sqrt{2}} \to \infty \quad \text{for } \alpha > 0. \end{aligned}$$

We know from proposition 4.1.1 that we may always find a twice differentiable square root of a non-negative smooth function, so that the eigenvalues λ are functions which are twice differentiable but not $C^{1,\alpha}$ for any $\alpha > 0$.

14.2.3. Lemma (Resolvent lemma). *If $A(t)$ is $C^{k,\alpha}$ for some $1 \le k \le \infty$ and $\alpha > 0$, then the resolvent $(t, z) \mapsto (A(t) - z)^{-1} \in L(H, H)$ is C^k on its natural domain. (By $C^{\infty,\alpha}$ we mean C^∞.)*

Proof. By definition the function $t \mapsto \langle A(t)u \mid v\rangle$ is of class $C^{k,\alpha}$ for each $u \in V$ and $u \in H$. By [**KM97**, 2.3] (extended from $C^{k,1}$ to $C^{k,\alpha}$ with essentially the same proof), the curve $t \mapsto A(t)u$ is of class $C^{k,\alpha}$ into H.

For each t we consider the norm $\|u\|_t^2 := \|u\|^2 + \|A(t)u\|^2$ on V. Since $A(t) = A^*(t)$ is closed, the pair $(V, \|.\|_t)$ is again a Hilbert space with inner product given by $\langle u \mid v\rangle_t := \langle u \mid v\rangle + \langle A(t)u \mid A(t)v\rangle$.

We claim that all these norms $\|.\|_t$ on V are locally uniformly equivalent.

For: The operator $A(t) : (V, \|.\|_s) \to H$ is bounded, since the graph of $A(t)$ is closed in $H \times H$, contained in $V \times H$ and thus also closed in $(V, \|.\|_s) \times H$. For fixed $u, v \in V$, the function $t \mapsto \langle u \mid v\rangle_t = \langle u \mid v\rangle + \langle A(t)u \mid A(t)v\rangle$ is $C^{k,\alpha}$, by the remark at the beginning of the proof. Thus it is also locally Lipschitz ($C^{0,1}$). By the multilinear uniform boundedness principle ([**KM97**, 5.8]), the mapping $t \mapsto \langle . \mid .\rangle_t$ is $C^{0,1}$ into the space of bounded bilinear forms on $(V, \|.\|_s)$ for each fixed s. By the exponential law ([**KM97**, 3.12]), the mapping $(t, u, v) \mapsto \langle u \mid v\rangle_t$ is $C^{0,1}$ from $\mathbb{R} \times (V, \|.\|_s) \times (V, \|.\|_s)$ to $\mathbb{R}$ for each s. Therefore, and by homogeneity in (u, v) the set $\{\|u\|_t : |t| \le K, \|u\|_s \le 1\}$ is bounded by some $L_{K,s}$ in $\mathbb{R}$. Thus $\|u\|_t \le L_{K,s}\|u\|_s$ for all $|t| \le K$, i.e., all norms $\|.\|_t$ are locally uniformly equivalent. This shows the assertion.

By [**FF89**, 5], and the linear uniform boundedness theorem we have that $t \mapsto A(t)$ is a $C^{k,\alpha}$ mapping $\mathbb{R} \to L(V, H)$, and thus C^k in the usual sense, again by [**FF89**, 5]. Alternatively, one may use [**KM97**, 2.3], extended from $C^{k,1}$ to $C^{k,\alpha}$, and [**KM97**, 5.18], which says that it suffices to test with linear maps which recognize bounded sets.

If for some $(t,z) \in \mathbb{R} \times \mathbb{C}$ the bounded operator $A(t) - z : V \to H$ is invertible, then this is true locally and $(t,z) \mapsto (A(t) - z)^{-1} : H \to V$ is C^k, by the chain rule, since inversion is smooth on the Banach space $L(V,H)$. □

Note that the statement of the resolvent lemma cannot be improved. In [**KM03**] there is given an example of a curve $A(t)$ of self-adjoint unbounded operators on ℓ^2 with compact resolvent and common domain V of definition, such that $t \mapsto \langle A(t)u \mid v\rangle$ is C^1 for all $u \in V$ and $v \in \ell^2$, but $t \mapsto A(t)$ is not differentiable at 0 into $L(V, \ell^2)$.

Since each $A(t)$ is Hermitian with compact resolvent the *global resolvent set* $\{(t,z) \in \mathbb{R} \times \mathbb{C} : (A(t) - z) : V \to H \text{ is invertible}\}$ is open and connected. Moreover, $(A(t) - z)^{-1} : H \to H$ is a compact operator for some (equivalently any) (t,z) if and only if the inclusion $i : V \to H$ is compact, since $i = (A(t) - z)^{-1} \circ (A(t) - z) : V \to H \to H$.

14.2.4. Let us return to the proof of theorem 14.2.2.

Claim (1). *Suppose that $A(t)$ is $C^{1,\alpha}$. Let z be an eigenvalue of $A(s)$ of multiplicity N. Then there exists an open box $(s-\delta, s+\delta) \times (z-\epsilon, z+\epsilon)$ and C^1 functions $\mu_1, \dots, \mu_N : (s-\delta, s+\delta) \to (z-\epsilon, z+\epsilon)$ which parameterize all eigenvalues λ with $|\lambda - z| < \epsilon$ of $A(t)$ for $|t-s| < \delta$ with correct multiplicities.*

Proof. Choose a simple closed smooth curve γ in the resolvent set of $A(s)$ for fixed s enclosing only z among all eigenvalues of $A(s)$. Since the global resolvent set is open, no eigenvalue of $A(t)$ lies on γ, for t near s. Since

$$t \mapsto -\frac{1}{2\pi i} \int_\gamma (A(t) - z)^{-1} dz =: P(t,\gamma) = P(t)$$

is a C^1 curve of projections (on the direct sum of all eigenspaces corresponding to eigenvalues in the interior of γ) with finite dimensional ranges, the ranks (i.e. dimensions of the ranges) must be constant: It is easy to see that the (finite) rank cannot fall locally, and it cannot increase, since the distance in $L(H,H)$ of $P(t)$ to the subset of operators of rank less or equal $N = \operatorname{rank} P(s)$ is continuous in t and is either 0 or 1. So for t near s, say $t \in I := (s-\delta, s+\delta)$, there are equally many eigenvalues in the interior of γ, and we may call them $\lambda_i(t)$ for $1 \le i \le N$ (repeated with multiplicities), so that each λ_i is continuous. The image of $t \mapsto P(t,\gamma)$, for t near s, describes a C^1 finite dimensional vector subbundle of $\mathbb{R} \times H \to \mathbb{R}$, since its rank is constant. For each t choose an orthogonal system of eigenvectors $v_j(t)$ of $A(t)$ corresponding to these $\lambda_j(t)$. They form a (not necessarily continuous) framing of this bundle. For any t near s and any sequence $(t_k)_k$ with $t_k \to t$ there is a subsequence again denoted by $(t_k)_k$ such that each $v_j(t_k) \to w_j(t)$, where the $w_i(t)$ form again an orthonormal system of eigenvectors of $A(t)$ for the sum $P(t)(H)$ of the eigenspaces of the $\lambda_i(t)$; here local triviality of the vector bundle is used. Consider

$$\frac{A(t) - \lambda_i(t)}{t_k - t} v_i(t_k) + \frac{A(t_k) - A(t)}{t_k - t} v_i(t_k) - \frac{\lambda_i(t_k) - \lambda_i(t)}{t_k - t} v_i(t_k) = 0. \tag{14.1}$$

For $t = s$ we take the inner product of (14.1) with each $w_j(s)$. Then the first summand vanishes, since all $\lambda_i(s)$ agree. If $k \to \infty$ we obtain that, for $i \ne j$, the $w_i(s)$ are a basis of eigenvectors of $P(s)A'(s)|_{P(s)(H)}$ with eigenvalues, for $i = j$,

$\lim_{k\to\infty} \frac{\lambda_i(t_k)-\lambda_i(s)}{t_k-s}$. Hence

$$\lim_{h\searrow 0} \frac{\lambda_i(s+h)-\lambda_i(s)}{h} = \rho_i,$$

where the ρ_i are the eigenvalues of $P(s)A'(s)|_{P(s)(H)}$ (with correct multiplicities), by lemma 11.4.1. So the right-sided derivative $\lambda_j^{(+)}(s)$ of each λ_j exists at s. Similarly the left-sided derivatives $\lambda_i^{(-)}(s)$ exist, and they form the same set of numbers with the correct multiplicities. Thus there exists a permutation σ of $\{1,\dots,N\}$ such that the functions

$$\nu_i(t) := \begin{cases} \lambda_i(t) & \text{for } t \le s \\ \lambda_{\sigma(i)}(t) & \text{for } t \ge s \end{cases} \tag{14.2}$$

parameterize all eigenvalues in the box by continuous functions which are differentiable at s.

For $t \neq s$, we take the inner product of (14.1) with each $w_i(t)$ to conclude that the right-sided derivative $\lambda_i^{(+)}(t)$ of λ_i at t exists and satisfies

$$\lambda_i^{(+)}(t) = \langle A'(t)w_i(t) \mid w_i(t)\rangle \tag{14.3}$$

for a unit eigenvector $w_i(t)$ of $A(t)$ with eigenvalue $\lambda_i(t)$.

Now we show claim 1 by induction on N. Let $t_1 \in I$ be such that not all $\lambda_i(t_1)$ agree. Then $\{1,\dots,N\}$ decomposes into the subsets $\{i : \lambda_i(t_1) = w\}$. For i and k in different subsets $\lambda_i(t) \neq \lambda_k(t)$ for all t in an open interval I_1 containing t_1. Thus, by induction, claim 1 holds on I_1.

Let $I_2 \subseteq I$ be an open interval containing only points t_1 as above. Let J be a maximal open subinterval of I_2 on which claim 1 holds. Assume that the right (say) endpoint b of J belongs to I_2. Then there is a C^1 parameterization of all N eigenvalues on an open interval I_b containing b by the argument above. Let $t_2 \in J \cap I_b$. Renumbering the C^1 parameterization to the right of t_2 suitably we may extend the C^1 parameterization beyond b. Hence, claim 1 holds on I_2.

Consider the closed set $E = \{t \in I : \lambda_1(t) = \dots = \lambda_N(t)\}$. Then $I\backslash E$ is open and so a disjoint union of open intervals on which there exists a C^1 parameterization μ_i of all eigenvalues. Consider first the set E' of all isolated points in E. Then $E' \cup (I\backslash E)$ is again open and thus a disjoint union of open intervals, and for each point $t \in E'$ we apply in turn the following arguments: Extending all μ_i by a single value at t we get a continuous extension near t. By (14.2), we may renumber the μ_i to the right of t in such a way that they fit together differentiably at t. The derivatives are also continuous at t: They have only finitely many cluster points for $t_k \to t$ by applying (14.3) to t_k and choosing a subsequence such that the $w_i(t_k)$ converge. Now we apply the arguments surrounding (14.1) with the $v_j(t_k)$ replaced by the $w_j(t_k)$ to conclude that (14.3) converges to $\rho_i(t) = \mu_i'(t)$. It follows that claim 1 holds on $E' \cup (I\backslash E)$.

We extend each μ_i to the whole of I by taking the single continuous function on $E\backslash E'$. Let $t \in E\backslash E'$. Then for the parameterization ν_i from (14.2) of all eigenvalues which is differentiable at t all derivatives $\nu_i'(t)$ agree, since t is a cluster point of E. Thus also $\mu_i'(t)$ exists and equals $\nu_i'(t)$. So all μ_i are differentiable on I.

To see that μ_i' is continuous at $t \in E\backslash E'$, let $t_k \to t$ be such that $\mu_i'(t_k)$ converges (to a cluster point or $\pm\infty$). By (14.3), we have $\mu_i'(t_k) = \langle A'(t_k)w_i(t_k) \mid w_i(t_k)\rangle$ for eigenvectors $w_i(t_k)$ of $A(t_k)$ with eigenvalue $\mu_i(t_k)$. Passing to a subsequence we may assume that the $w_i(t_k)$ converge to an orthonormal basis of eigenvectors of $A(t)$. Hence $\langle A'(t_k)w_i(t_k) \mid w_i(t_k)\rangle$ converges to some of the equal eigenvalues ρ_i of $P(t)A'(t)|_{P(t)(H)}$ which also equal the $\nu_i'(t)$.

This completes the proof of claim 1. $\square$

14.2.5.

Claim (2). *If $A(t)$ is smooth and no two of the generically different eigenvalues meet of infinite order, then the functions $\mu_1, \dots, \mu_N$ in claim 1 may be chosen smoothly.*

Proof. By replacing $A(s)$ by $A(s) - z_0$ if necessary we may assume that 0 is not an eigenvalue of $A(s)$. Let γ be a simple closed curve in the resolvent set of $A(s)$ for fixed s. Similarly as above we conclude that

$$t \mapsto -\frac{1}{2\pi i}\int_\gamma (A(t)-z)^{-1}dz =: P(t,\gamma) = P(t)$$

is a smooth curve of projections with finite dimensional ranges and constant ranks, and that for t near s, there are equally many eigenvalues $\mu_1(t), \dots, \mu_N(t)$ in the interior of γ. We denote by $v_i(t)$ a corresponding system of eigenvectors of $A(t)$. By the residue theorem we have

$$\sum_{i=1}^N \mu_i(t)^p v_i(t)\langle v_i(t) \mid \quad \rangle = -\frac{1}{2\pi i}\int_\gamma z^p (A(t)-z)^{-1}dz$$

which is smooth in t near s, as a curve of operators in $L(H,H)$ of rank N, since 0 is not an eigenvalue.

We assert that, for a smooth curve $t \mapsto T(t) \in L(H,H)$ of operators of rank N in a Hilbert space H such that $T(0)T(0)(H) = T(0)(H)$, also $t \mapsto \operatorname{trace}(T(t))$ is smooth near 0.

Namely: Let $F := T(0)(H)$. Then $T(t) = (T_1(t), T_2(t)) : H \to F \oplus F^\perp$ and the image of $T(t)$ is the space

$$\begin{aligned}
T(t)(H) &= \{(T_1(t)(x), T_2(t)(x)) : x \in H\}\\
&= \{(T_1(t)(x), T_2(t)(x)) : x \in F\} \quad \text{for } t \text{ near } 0\\
&= \{(y, S(t)(y)) : y \in F\}, \quad \text{where } S(t) := T_2(t) \circ (T_1(t)|_F)^{-1}.
\end{aligned}$$

Note that $S(t) : F \to F^\perp$ is smooth in t by finite dimensional inversion for $T_1(t)|_F : F \to F$. Now

$$\begin{aligned}
\operatorname{trace}(T(t)) &= \operatorname{trace}\left(\begin{pmatrix} 1 & 0 \\ -S(t) & 1\end{pmatrix}\begin{pmatrix} T_1(t)|_F & T_1(t)|_{F^\perp} \\ T_2(t)|_F & T_2(t)|_{F^\perp}\end{pmatrix}\begin{pmatrix} 1 & 0 \\ S(t) & 1\end{pmatrix}\right)\\
&= \operatorname{trace}\left(\begin{pmatrix} T_1(t)|_F & T_1(t)|_{F^\perp} \\ 0 & -S(t)T_1(t)|_{F^\perp} + T_2(t)|_{F^\perp}\end{pmatrix}\begin{pmatrix} 1 & 0 \\ S(t) & 1\end{pmatrix}\right)\\
&= \operatorname{trace}\left(\begin{pmatrix} T_1(t)|_F & T_1(t)|_{F^\perp} \\ 0 & 0\end{pmatrix}\begin{pmatrix} 1 & 0 \\ S(t) & 1\end{pmatrix}\right) \quad \text{since rank} = N\\
&= \operatorname{trace}\begin{pmatrix} T_1(t)|_F + (T_1(t)|_{F^\perp})S(t) & T_1(t)|_{F^\perp} \\ 0 & 0\end{pmatrix}\\
&= \operatorname{trace}\left(T_1(t)|_F + (T_1(t)|_{F^\perp})S(t) : F \to F\right),
\end{aligned}$$

which is visibly smooth, since F is finite dimensional. Hence the assertion.

Now we can conclude that

$$\sum_{i=1}^N \mu_i(t)^p = -\frac{1}{2\pi i}\operatorname{trace}\int_\gamma z^p(A(t)-z)^{-1}dz$$

is smooth for t near s.

It follows that the Newton polynomial mapping $s^N(\mu_1(t), \dots, \mu_N(t))$ is smooth, so also the elementary symmetric function mapping $\sigma^N(\mu_1(t), \dots, \mu_N(t))$ is smooth, and thus $\{\mu_i(t) : 1 \le i \le N\}$ is the set of roots of a polynomial with smooth coefficients, compare with section 3.1. By theorem 4.3.1, there is a smooth arrangement of these roots. □

14.2.6.

CLAIM (3). *Let I be a compact interval. Let $t \mapsto \lambda_i(t)$ be a differentiable eigenvalue of $A(t)$, defined on some subinterval of I. Then*

$$|\lambda_i(t_1) - \lambda_i(t_2)| \le (1 + |\lambda_i(t_2)|)(e^{a|t_1 - t_2|} - 1)$$

holds for a positive constant a depending only on I.

PROOF. From (14.3) we conclude, where $V_t = (V, \|.\|_t)$,

$$\begin{aligned}
|\lambda_i'(t)| &\le \|A'(t)\|_{L(V_t,H)} \|w_i(t)\|_{V_t} \|w_i(t)\|_H \\
&= \|A'(t)\|_{L(V_t,H)} \sqrt{\|w_i(t)\|_H^2 + \|A(t)w_i(t)\|_H^2} \\
&= \|A'(t)\|_{L(V_t,H)} \sqrt{1 + \lambda_i(t)^2} \\
&\le C + C|\lambda_i(t)|,
\end{aligned}$$

for a constant C, since all norms $\|.\|_t$ are locally in t uniformly equivalent, as seen in 14.2.3. By Gronwall's lemma (e.g. [**Die60**]) this implies claim 3. □

14.2.7. LEMMA. *Suppose that $\lambda_1, \dots, \lambda_N$ are real-valued C^1 (twice differentiable) functions defined on an interval I, and that $\mu_1, \dots, \mu_k$ for $k \le N$ are also C^1 (twice differentiable) functions on I such that $|\{j : \mu_j(t) = z\}| \le |\{i : \lambda_i(t) = z\}|$ for all $t \in I$ and $z \in \mathbb{R}$. Then there exist C^1 (twice differentiable) functions $\mu_{k+1}, \dots, \mu_N$ on I such that for all $t \in I$ and $z \in \mathbb{R}$ we have*

$$|\{j : 1 \le j \le N, \mu_j(t) = z\}| = |\{i : \lambda_i(t) = z\}|.$$

PROOF. Let us treat the case C^1 and indicate the necessary changes for the twice differentiable case.

We use induction on N. Assume that the statement is true if the number of functions is less than N.

First suppose that for given $t_1 \in I$ not all $\lambda_i(t_1)$ agree. Then for $i \in \{k : \lambda_1(t_1) = \lambda_k(t_1)\} \not\ni j$ we have $\lambda_i(t) \ne \lambda_j(t)$ for all t in an open interval I_1 containing t_1, and similarly for the μ_j. Thus, by induction hypothesis, for both groups the statement holds on I_1.

Now suppose that for no point t in I we have $\lambda_1(t) = \dots = \lambda_N(t)$. Let I_1 be a maximal open subinterval of I for which the statement is true with functions $\mu^1_{k+1}, \dots, \mu^1_N$. Assume for contradiction that the right (say) endpoint b of I_1 is an interior point of I. By the previous paragraph the statement holds for an open neighborhood I_2 of b, with functions $\mu^2_{k+1}, \dots, \mu^2_N$. Let $t_0 \in I_1 \cap I_2$. We may continue each solution μ^1_i on $\{t \in I_1 : t \le t_0\}$ by a suitable solution $\mu^2_{\sigma(i)}$ on $\{t \in I_2 : t \ge t_0\}$ for a suitable permutation σ. Namely: Let $t_m \nearrow t_0$. For every m there exists a permutation σ of $\{1, \dots, N\}$ such that $\mu^2_{\sigma(i)}(t_m) = \mu^1_i(t_m)$ for all i. By passing to a subsequence, again denoted by $(t_m)_m$, we may assume that the permutation does not depend on m. Passing to a subsequence again we can also assume that $(\mu^2_{\sigma(i)})'(t_m) = (\mu^1_i)'(t_m)$ (and in the twice differentiable case, again for a subsequence, we get $(\mu^2_{\sigma(i)})''(t_m) = (\mu^1_i)''(t_m)$ for all i and all m. So we may paste $\mu^2_{\sigma(i)}(t)$ for $t \ge t_0$ with $\mu^1_i(t)$ for $t < t_0$ to obtain a C^1 (twice differentiable) parameterization on an interval larger than I_1, a contradiction.

In the general case, consider the closed set $E = \{t \in I : \lambda_1(t) = \ldots = \lambda_N(t)\}$. Then $I \backslash E$ is open, thus a disjoint union of open intervals. By the last paragraph the result holds on each of these open intervals. Let E' denote the set of all isolated points of E. Then $E' \cup (I \backslash E)$ is open and thus a union of open intervals. For each point $t \in E'$ we can renumber the μ_i to the right of t in such a way that they fit together C^1 (twice differentiable) at t. Hence the result holds on $E' \cup (I \backslash E)$.

Finally extend each μ_i to the whole of I by taking the single continuous function on $E\backslash E'$. Let $t \in E\backslash E'$. Then all $\lambda_i'(t) =: \lambda'(t)$ agree, since t is a cluster point of E (and all $\lambda_i''(t) =: \lambda''(t)$ agree by considering second order difference quotients on points of E). Thus μ_i is (twice) differentiable at t with $\mu_i'(t) = \lambda'(t)$ (and $\mu_i''(t) = \lambda''(t)$). In the C^1 case we still have to check that μ_i' is continuous at $t \in E\backslash E'$: Let $t_n \to t$, then $\mu_i'(t_n) = \lambda'_{\sigma_n(i)}(t_n) \to \lambda'(t) = \mu_i'(t)$. □

14.2.8. Proof of theorem 14.2.2. To (1): By claim 3 no eigenvalue can go off to infinity in finite time, since it may increase at most exponentially.

Let us first number all eigenvalues of $A(0)$ increasingly. We claim that for one eigenvalue (say $\lambda_0(0)$) there exists a smooth extension to all of $\mathbb{R}$. For: The set of all $t \in \mathbb{R}$ with a smooth extension of λ_0 on the segment from 0 to t is open and closed. Open follows from claim 2. If this interval does not reach infinity, from claim 3 it follows that $(t, \lambda_0(t))$ has a cluster point (s, x) at the end s. Clearly, x is an eigenvalue of $A(s)$, and by claim 2, the eigenvalues passing through (s, x) can be arranged smoothly. Thus $\lambda_0(t)$ converges to x and can be extended differentiably beyond s.

By the same argument we can extend iteratively all eigenvalues smoothly to all $t \in \mathbb{R}$: If it meets an already chosen one, the proof of theorem 4.3.1 shows that we may pass through it coherently.

Let us consider eigenvalues $\{\lambda_i(t) : 1 \le i \le N\}$ contained in the interior of a smooth curve γ for t in an open interval I. Then $V_t := P(t, \gamma)(H)$ is the fiber of a smooth vector bundle of dimension N over I. We choose a smooth framing of this bundle, and use then the proof of theorem 14.1.2 to choose smooth vector subbundles whose fibers over t are the eigenspaces of the eigenvalues with their generic multiplicity. Similarly as in the proof of theorem 14.1.2 we then get global vector subbundles with fibers the eigenspaces of the eigenvalues with their generic multiplicity, and finally smooth eigenvectors for all eigenvalues.

To (3): We number all eigenvalues of $A(0)$ increasingly. Consider families of C^1 functions $(\mu_i)_{i\in\alpha}$ indexed by ordinals α, defined on open intervals I_i containing some fixed t_0, which parameterize eigenvalues. The set of these sequences is partially ordered by inclusion of ordinals and then by restriction of the component functions. Obviously, for each increasing chain of such sequences the union is again such a sequence. By Zorn's lemma, there exists a maximal family (μ_i).

We claim that for each maximal family each component function μ_i is globally defined: If not, let $b < \infty$ be the right (say) boundary point of I_i. By claim 3, the limit $\lim_{t\nearrow b} \mu_i(t) =: z$ exists. By claim 1, there is a box $(b-\delta, b+\delta)\times(z-\epsilon, z+\epsilon)$ such that all eigenvalues λ of $A(t)$ with $|\lambda - z| < \epsilon$ for $|t-b| < \delta$ are parameterized by C^1 functions $\lambda_i : (b-\delta, b+\delta) \to (z-\epsilon, z+\epsilon)$ (with multiplicity). Consider the μ_j hitting this box (at the vertical boundaries only). The endpoints of the corresponding intervals I_j give a partition of $(b-\delta, b+\delta)$ into finitely many subintervals. Applying lemma 14.2.7 on each subinterval and gluing at the ends of the subintervals in C^1 fashion, using (14.2), we obtain an extension of at least μ_i. So the family was not maximal.

Finally we assert that any family (μ_i) parameterizes all eigenvalues of $A(t)$ with right multiplicities, for all $t \in \mathbb{R}$. If not, then there is an eigenvalue z of $A(t_0)$ with $|\{i : \mu_i(t_0) = z\}|$ less than the multiplicity of z. By claim 1 and lemma 14.2.7, we can then conclude again that the sequence was not maximal.

To (2): We shall show the following statement which implies (2):

If the multiplicity of an eigenvalue never exceeds n, and if the resolvent map $(A(t) - z)^{-1}$ is C^{3n} into $L(H, H)$ in t and z jointly, then the eigenvalues of $A(t)$ may be parameterized twice differentiable in t.

Choose a simple closed smooth curve γ in the resolvent set of $A(s)$ for fixed s enclosing only z among all eigenvalues of $A(s)$. As in the proof of claim 2 we can conclude that, for t near s, there are equally many eigenvalues $\mu_1(t), \dots, \mu_N(t)$ in the interior of γ. If we denote by $v_i(t)$ a corresponding system of eigenvectors of $A(t)$, then

$$\sum_{i=1}^{N} \mu_i(t)^p v_i(t) \langle v_i(t) \mid \quad \rangle = -\frac{1}{2\pi i} \int_\gamma z^p (A(t) - z)^{-1} dz$$

is C^{3n} in t near s, as a curve of operators in $L(H, H)$ of rank N. Moreover, for a C^{3n} curve $t \mapsto T(t) \in L(H, H)$ of operators of rank N in a Hilbert space H such that $T(0)T(0)(H) = T(0)(H)$, also $t \mapsto \operatorname{trace}(T(t))$ is C^{3n} near 0, by 14.2.5.

It follows that the Newton polynomials

$$s_p(t) := \sum_{i=1}^{N} \mu_i(t)^p = -\frac{1}{2\pi i} \operatorname{trace} \int_\gamma z^p (A(t) - z)^{-1} dz$$

are C^{3n} for t near s. Hence also the elementary symmetric functions

$$\sigma_p(t) := \sum_{1 \le i_1 < \dots < i_p \le N} \mu_{i_1}(t) \cdots \mu_{i_p}(t)$$

are C^{3n}. Therefore, $\{\mu_i(t) : 1 \le i \le N\}$ is the set of roots of a polynomial of degree $N \le n$ with C^{3n} coefficients. By theorem 7.1.1, there exists an arrangement of these roots such that they become twice differentiable.

The rest of the proof is analogous to the end of the proof of (3).

CHAPTER 15

Hyperbolic polynomials in a more general sense

15.1. Hyperbolic polynomials in many variables

We present a more general notion of hyperbolic polynomials originally motivated by partial differential equations literature, [**Går51**], [**Hör94**]. The section is mainly based on [**BGLS01**].

Let V be a finite dimensional real vector space.

15.1.1. DEFINITION. Suppose that P is a homogeneous polynomial of degree m on V and $N \in V$ with $P(N) \neq 0$. Then P is *hyperbolic with respect to* N, if the polynomial $\tau \mapsto P(\xi + \tau N)$, where τ is a scalar, has only real roots, for every $\xi \in V$.

15.1.2. DEFINITION. Suppose P is hyperbolic with respect to $N \in V$ of degree m. Then for every $\xi \in V$, we can write

$$P(\xi + \tau N) = P(N) \prod_{i=1}^{m} (\tau + \lambda_i(\xi))$$

and assume without loss that $\lambda_1(\xi) \geq \lambda_2(\xi) \geq \cdots \geq \lambda_m(\xi)$. The corresponding map $\xi \mapsto (\lambda_1(\xi), \cdots, \lambda_m(\xi))$ is denoted by λ and called the *characteristic map (with respect to P and N)*. We say that $\lambda_i(\xi)$ is the *i-th largest characteristic root (with respect to P and N)* and define *the sum of the k largest characteristic roots* by $\theta_k := \sum_{i=1}^{k} \lambda_i$, for $1 \leq k \leq m$. The function θ_m is called the *trace*.

Let us assume for this section that P is a homogeneous polynomial of degree m which is hyperbolic with respect to N with characteristic map λ.

15.1.3. EXAMPLE (Hermitian matrices). Let V be the real vector space of the $m \times m$ Hermitian matrices and $P :=$ det. Then P is hyperbolic of degree m with respect to the unit matrix I and λ maps $\xi \in V$ to its eigenvalues, arranged decreasingly. Here θ_m is the ordinary trace.

15.1.4. A simple way to generate new hyperbolic polynomials is differentiation: *If $m > 1$, then $Q(\xi) := \frac{d}{d\tau}\big|_{\tau=0} P(\xi + \tau N)$ is hyperbolic with respect to N*, which is basically a consequence of Rolle's theorem.

15.1.5. The following property of the characteristic roots is proved in [**Går59**]: *For all $r, s \in \mathbb{R}$ and every $1 \leq i \leq m$*

$$\lambda_i(r\xi + sN) = \begin{cases} r\lambda_i(\xi) + s, & \textit{if } r \geq 0; \\ r\lambda_{m+1-i}(\xi) + s, & \textit{otherwise.} \end{cases}$$

It follows that the characteristic map λ is positively homogeneous and continuous. The following theorem is due to L. Gårding.

15.1.6. THEOREM ([**Går51**], [**Går59**]). *The largest root λ_1 is a* sublinear *function, i.e., positively homogeneous and convex.*

15.1.7. DEFINITION. The *hyperbolicity cone $C(N)$ of P with respect to N*, is the set $C(N) = \{\xi \in V : P(\xi - \tau N) \neq 0 \text{ for all } \tau \leq 0\}$.

We have $C(N) = \{\xi \in V : \lambda_m(\xi) > 0\}$. So $C(N)$ is an open convex cone that contains N with closure $\overline{C(N)} = \{\xi \in V : \lambda_m(\xi) \geq 0\}$. If $M \in C(N)$, then P is hyperbolic with respect to M and $C(M) = C(N)$; see [**Går51**], [**Går59**].

Definition. P is *complete* if $\{\xi \in V : \lambda(\xi) = 0\} = \{0\}$.

We have $\{\xi \in V : \lambda(\xi) = 0\} = \overline{C(N)} \cap (-\overline{C(N)})$, [**Går51**], [**Går59**].

Example. We continue example 15.1.3. The hyperbolicity cone of $P = \det$ with respect to the unit matrix I is the set of all positive definite matrices. The polynomial $P = \det$ is complete, since every non-zero Hermitian matrix has at least one non-zero eigenvalue.

15.1.8. Example (Elementary symmetric functions). Let σ_i denote the i-th elementary symmetric function, see e.g. section 3.1. For every $\xi \in V$ and all $\tau \in \mathbb{R}$,

$$P(\xi + \tau N) = P(N) \prod_{i=1}^{m} (\tau + \lambda_i(\xi)) = P(N) \sum_{i=0}^{m} \sigma_i(\lambda(\xi)) \tau^{m-i},$$

and for any $0 \leq i \leq m$,

$$P(N)\sigma_i(\lambda(\xi)) = \frac{1}{(m-i)!} d^{m-i} P(\xi)[N, N, \ldots, N].$$

If $1 \leq i \leq m$, then $\sigma_i \circ \lambda$ is hyperbolic with respect to N of degree i. In particular, the trace $\theta_m = \sigma_1 \circ \lambda$ is a homogeneous hyperbolic polynomial of degree 1 and hence linear. Moreover, the elementary symmetric polynomials themselves are hyperbolic: Let $V = \mathbb{R}^m$ and $N = (1, 1, \ldots, 1)$. It is evident that σ_m is homogeneous of degree m and hyperbolic with respect to N with corresponding characteristic map $\lambda(\xi) = \xi_\downarrow$, where $\xi_\downarrow$ is the vector ξ with its coordinates arranged decreasingly. By the above, σ_i, for $1 \leq i \leq m-1$, is homogeneous of degree i and hyperbolic with respect to N.

15.1.9. Constructing new hyperbolic polynomials. The class of hyperbolic polynomials form a rich class. There are various ways of constructing new hyperbolic polynomials from old ones, e.g. differentiation as seen in 15.1.4. Here is another construction due to H.H. Bauschke, O. Güler, A.S. Lewis, and H.S. Sendov.

Theorem ([**BGLS01**]). *Suppose Q is a homogeneous symmetric polynomial of degree n on $\mathbb{R}^m$, hyperbolic with respect to $M = (1, 1, \ldots, 1)$, with characteristic map μ. Then $Q \circ \lambda$ is a homogeneous polynomial of degree n which is hyperbolic with respect to N and its characteristic map is $\mu \circ \lambda$.*

This yields a generalization of Gårding's theorem 15.1.6:

Corollary. *We have :*

(1) *For each $1 \leq k \leq m$, the function θ_k is sublinear.*
(2) *Each characteristic root is globally Lipschitz.*
(3) *The function $w^\top \lambda$ is sublinear for any $w \in \mathbb{R}^m_\downarrow = \{u \in \mathbb{R}^m : u_1 \geq \cdots \geq u_m\}$.*
(4) *For $\xi, \eta \in V$ we have $\|\lambda(\xi + \eta)\| \leq \|\lambda(\xi) + \lambda(\eta)\|$. Equality holds if and only if $\lambda(\xi + \eta) = \lambda(\xi) + \lambda(\eta)$.*

Proof. To (1): Fix k and set

$$Q(u) = \prod_{1 \leq i_1 < \cdots < i_k \leq m} \sum_{l=1}^{k} u_{i_l}.$$

Then Q is a homogeneous symmetric polynomial on $\mathbb{R}^m$ of degree $\binom{m}{k}$, hyperbolic with respect to $(1, \ldots, 1)$, and its characteristic roots are given by

$$\left\{ \frac{1}{k} \sum_{l=1}^{k} u_{i_l} : 1 \leq i_1 < \cdots < i_k \leq m \right\}.$$

By theorem 15.1.9, the largest characteristic root of $Q \circ \lambda$ is equal to $\frac{1}{k}\theta_k(\xi)$. Now theorem 15.1.6 yields the sublinearity of θ_k.

To (2): It is well-known that every sublinear finite function is globally Lipschitz (e.g. [**Roc70**]), in particular, each θ_k, by (1). Thus $\lambda_1 = \theta_1$ and $\lambda_k = \theta_k - \theta_{k-1}$, for $2 \leq k \leq m$, are Lipschitz.

To (3): We may write $w^\top \lambda = \sum_{i=1}^m w_i \lambda_i = w_m \theta_m + \sum_{i=1}^{m-1}(w_i - w_{i+1})\theta_i$. Then (3) follows from (1).

To (4): Let $w = \lambda(\xi + \eta) \in \mathbb{R}^m_\downarrow$. Then, using (3) and the Cauchy-Schwarz inequality in $\mathbb{R}^m$, we estimate

$$\begin{aligned} \|\lambda(\xi + \eta)\|^2 &= w^\top \lambda(\xi + \eta) \leq w^\top (\lambda(\xi) + \lambda(\eta)) \\ &\leq \|w\| \|\lambda(\xi) + \lambda(\eta)\| = \|\lambda(\xi + \eta)\| \|\lambda(\xi) + \lambda(\eta)\|. \end{aligned}$$

The condition for equality follows from the condition for equality in the Cauchy-Schwarz inequality. □

15.1.10. The characteristic map λ induces a natural norm on V, if P is complete:

$$\begin{aligned} \|\cdot\| : V &\longrightarrow [0, +\infty) \\ \xi &\longmapsto \|\lambda(\xi)\|. \end{aligned}$$

This follows from 15.1.5 and corollary 15.1.9(4). For this norm we obtain a sharpened Cauchy-Schwarz inequality:

THEOREM. *Suppose P is complete. Then*

$$\langle \xi \mid \eta \rangle \leq \langle \lambda(\xi) \mid \lambda(\eta) \rangle \leq \|\xi\| \|\eta\|,$$

for all $\xi, \eta \in V$.

PROOF. The Cauchy-Schwarz inequality in $\mathbb{R}^m$ and corollary 15.1.9(4) imply

$$\begin{aligned} 2\langle \lambda(\xi) \mid \lambda(\eta) \rangle &\geq \|\lambda(\xi + \eta)\|^2 - \|\lambda(\xi)\|^2 - \|\lambda(\eta)\|^2 \\ &= \|\xi + \eta\|^2 - \|\xi\|^2 - \|\eta\|^2 = 2\langle \xi \mid \eta \rangle, \end{aligned}$$

as required. □

EXAMPLE. The induced inner product on the Hermitian matrices is precisely what one would expect: $\langle \xi \mid \eta \rangle = \mathrm{trace}(\xi\eta)$. The sharpening of the Cauchy-Schwarz inequality is essentially due to von Neumann, see [**Lew96**, Theorem 2.2].

EXAMPLE. Consider the vector space $V = \mathbb{R}^n$, the polynomial $P(\xi) = \prod_{i=1}^n \xi_i$, and the direction $N = (1, 1, \ldots, 1)$. Then N is hyperbolic and complete with characteristic map $\lambda(\xi) = \xi_\downarrow$. The induced norm in V is just the standard Euclidean norm in $\mathbb{R}^n$. The sharpened Cauchy-Schwarz inequality reduces to the well-known Hardy-Littlewood-Pólya inequality ([**HLP52**, Chapter X]) $\xi^\top \eta \leq \xi_\downarrow^\top \eta_\downarrow$. Equality holds if and only if the vectors ξ and η can be simultaneously ordered with the same permutation.

More examples and more results may be found in [**BGLS01**].

15.1.11. A related, but a bit more general notion of hyperbolic polynomials is the following.

Definition. A polynomial P of degree m in n variables $\xi = (\xi_1, \ldots, \xi_n)$ and with principal part P_m is said to be *hyperbolic with respect to a real vector N* if $P_m(N) \neq 0$ and for some τ_0

$$P(\xi + \tau N) \neq 0 \qquad \text{if } \xi \in \mathbb{R}^n \text{ and } \mathrm{Im}(\tau) < \tau_0.$$

A hyperbolic polynomial P is said to be *strictly hyperbolic* if the roots of $\tau \mapsto P_m(\xi + \tau N)$ are simple for any real ξ which is not proportional to N. A hyperbolic polynomial P is called *normalized* if $P_m(N) = 1$.

For homogeneous polynomials P this notion of hyperbolicity coincides with the notion introduced in definition 15.1.1: If $P(\xi + \tau N) = 0$, then $P(\sigma\xi + \sigma\tau N) = 0$ for every real σ, by homogeneity. If P is hyperbolic with respect to N, we have $\sigma\mathrm{Im}(\tau) \geq \tau_0$ for every real σ. Hence $\mathrm{Im}(\tau) = 0$. Conversely, if $P(\xi + \tau N) = 0$ has only real roots, then $P(\xi + i\tau N) \neq 0$ for $\tau \neq 0$, so P is hyperbolic.

If a polynomial P is hyperbolic with respect to N so is its principal part P_m: The roots of $P(\xi + i\tau N) = 0$ are located in the half plane $\mathrm{Re}(\tau) \geq \tau_0$. We have

$$P_m(\xi + i\tau N) = \lim_{\sigma \to +\infty} \sigma^{-m} P(\sigma\xi + i\sigma\tau N).$$

Since the roots of $P(\sigma\xi + i\sigma\tau N) = 0$ lie in the half plane $\sigma\mathrm{Re}(\tau) \geq \tau_0$, we may conclude that the roots of $P_m(\xi + i\tau N) = 0$ all lie in the half plane $\mathrm{Re}(\tau) \geq 0$. That shows the assertion.

The importance of this notion in the theory of PDE's is the following: The Cauchy problem for the differential operator $P(D)$ with data on a non-characteristic hyperplane $\langle x \mid N \rangle = 0$ cannot be solved in general unless P is hyperbolic with respect to N. If $P(\xi)$ is a polynomial in n variables $\xi_1, \ldots, \xi_n$, with complex coefficients, then a differential operator $P(D)$ is defined by replacing ξ_j by $D_{x_j} = -i\frac{\partial}{\partial x_j}$. That the hyperplane $\langle x \mid N \rangle = 0$ is *non-characteristic* means that $P_m(N) \neq 0$. See [**Hör83b**, Chapter XII].

The following theorem due to W. Nuij deals with topological properties of the space of polynomials of given degree which are hyperbolic with respect to a given vector.

Theorem ([**Nui68**]). *Let us define a topology in a space of polynomials of given degree by using the Euclidean norm for the coefficients. In the following statements the polynomials are understood to have a fixed degree and to be hyperbolic with respect to a fixed vector. We have:*

(1) *The space of strictly hyperbolic (homogeneous) polynomials is open.*
(2) *Every hyperbolic (homogeneous) polynomial is the limit of strictly hyperbolic (homogeneous) polynomials.*
(3) *The space of (strictly) hyperbolic polynomials is contractible to the space of (strictly) hyperbolic homogeneous polynomials.*
(4) *The space of normalized (strictly) hyperbolic (homogeneous) polynomials is connected and simply connected.*

The proof of this theorem uses the splitting operator $P \mapsto P + sP'$ that reduces the multiplicity of the multiple roots of P. We made use of that splitting operator in chapter 6.

For more on the properties of hyperbolic polynomials in the sense of definition 15.1.11 we refer to [**Går51**], [**Går59**], [**Hör63**], [**Hör83a**], [**Hör83b**].

CHAPTER 16

Related lifting problems

16.1. Lifting diffeomorphisms of orbit spaces

16.1.1. Schwarz's isotopy lifting theorem. Let G be a compact Lie group and $\rho : G \to \mathrm{O}(V)$ an orthogonal representation in a real finite dimensional Euclidean vector space V. Let $\sigma_1, \dots, \sigma_n$ be a system of homogeneous generators of $\mathbb{R}[V]^G$ and $\sigma = (\sigma_1, \dots, \sigma_n)$ the orbit map. We identify $\sigma(V)$ with V/G and σ with the orbit projection $V \to V/G$. The orbit space V/G has a real analytic or smooth structure defined by the sheaf of G-invariant functions (see 8.2.4). Then the notions of diffeomorphism, isotopy, and etc. of orbit spaces have their usual categorical meaning.

Let $f : V/G \to V/G$ be a diffeomorphism. A diffeomorphism $F : V \to V$ is called a lift of f if $\sigma \circ F = f \circ \sigma$. We will investigate the problem of lifting for diffeomorphisms of the orbit space V/G. This section is based on [**Los01**].

If f belongs to the identity component of $\mathrm{Diff}(V/G)$, the group of diffeomorphisms of V/G, the above problem of lifting is solved by Schwarz's isotopy lifting theorem.

Theorem ([**Sch80**]). *Let M be a smooth G-manifold, G a compact Lie group. Suppose that $\bar{F}$ is a smooth isotopy of M/G starting at the identity. Then there is a smooth equivariant isotopy F of M starting at the identity and inducing $\bar{F}$.*

The real analytic version of theorem 16.1.1 is true as well.

16.1.2. Quasi-linear diffeomorphisms. Let $f : \sigma(V) \to \sigma(V)$ be a diffeomorphism. Since the sheaf $C^\infty(V/G)$ is uniquely defined by the generators $\sigma_1, \dots, \sigma_n$, by Schwarz's theorem 8.2.5, f is uniquely defined by the images of these generators as global sections of $C^\infty(V/G)$. We may consider f as the restriction to $\sigma(V)$ of some smooth map $\mathbb{R}^n \to \mathbb{R}^n$. Then we may write $f = (f_1, \dots, f_n)$, where $f_i \in C^\infty(\mathbb{R}^n)$. A diffeomorphism f is called *quasi-linear* if it can be written in the form $f = (f_1, \dots, f_n)$, where f_i are homogeneous polynomials in graded indeterminates y_j of degree d_j and $\deg f_i = d_i$.

Note that in the real analytic case the respective statements follow from Luna's analog of Schwarz's theorem, see [**Lun76**].

Suppose V^G is the subspace of G-invariant points of V and V_1 is the orthogonal complement of V^G in V. Then $\mathbb{R}[V]^G = \mathbb{R}[V^G] \otimes \mathbb{R}[V_1]^G$, $V/G = V^G \times (V_1/G)$, and we may choose generators $\sigma_1, \dots, \sigma_n$ so that $\sigma_1, \dots, \sigma_p$ are generators of $\mathbb{R}[V_1]^G$, where $\sigma_1(v) = \langle v \mid v \rangle$ and $\sigma_{p+1}, \dots, \sigma_n$ are generators of $\mathbb{R}[V^G]$. Then $d_1 = 2$, $d_{p+1} = \cdots = d_n = 1$ and $\sigma(0) = 0 \in \mathbb{R}^n$.

Theorem. *Let f be a diffeomorphism of $\sigma(V)$. Then there is a piecewise smooth isotopy f_t $(0 \le t \le 1)$ such that $f_1 = f$ and f_0 is a quasi-linear diffeomorphism of $\sigma(V)$.*

Proof. It is easily seen that $f(V^G \times \{0\}) = V^G \times \{0\}$. Therefore, $f(0) = v_0 \in V^G$. Let $f(x) = (f'(x), f''(x))$, where $f'(x) \in V^G$ and $f''(x) \in (V_1/G)$. Then $\bar{f} = (f' - v_0, f'')$ is a diffeomorphism of V/G and $\bar{f}(0) = 0$. It is obvious that there is a smooth isotopy joining f and $\bar{f}$. So we may reduce the proof to the case when $f(0) = 0$.

Let $f = (f_1, \ldots, f_n)$, where $f_i \in C^\infty(\mathbb{R}^n)$. Put $\deg y_i = d_i$. Let $T(f_i)$ be a part of the Taylor series of f_i of degree less than d_i. For $v \in V$, put $c(t) = f \circ \sigma(tv)$. By lemma 10.1.3, $m(c_i) \geq d_i$, and, since f is a diffeomorphism, $m(c_i) = d_i$. This is possible only if $T(f_i)$ vanishes on $\sigma(V)$. Then in the above presentation of f we can replace f_i with $f_i - T(f_i)$, i.e., we can suppose that all terms of the Taylor series of f_i have degrees more or equal to d_i.

Let, for each i, p_i be the part of the Taylor series of f_i homogeneous of degree d_i. Put $p = (p_1, \ldots, p_n)$, $ty = (t^{d_1}y_1, \ldots, t^{d_n}y_n)$, $f_{t,i}(y) = t^{-d_i}f_i(ty)$, and $h_t = (f_{t,i})$ for $t \neq 0$. It is clear that h_t is a diffeomorphism of $\sigma(V)$. For $y \in \sigma(V)$, $h_t(y)$ converges (uniformly for y in a bounded set) to $p(y)$ as $t \to 0$.

Put $g = f^{-1} = (g_i)$ and $g_t(y) = (t^{-d_i}g_i(ty))$. We can assume that all terms of the Taylor series of g_i have degrees more or equal to d_i and then $\lim_{t\to 0} g_t$ exists. It is clear that $g_t = h_t^{-1}$ for all $t \in [0, 1]$. Therefore, h_t is a smooth isotopy, $h_1 = f$, and $h_0 = p$ is a quasi-linear diffeomorphism. □

Theorems 16.1.1 and 16.1.2 imply the following

Corollary. *Let $\rho : G \to \mathrm{O}(V)$ be an orthogonal representation of a compact Lie group G. Suppose each quasi-linear diffeomorphism of V/G has a real analytic lift. Then:*

(1) *Each smooth diffeomorphism of V/G has a smooth lift.*
(2) *Each real analytic diffeomorphism of V/G has a real analytic lift.* □

16.1.3. Let us suppose now that the group G is finite.

Consider the complexification $V^{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C}$ of the vector space V, the induced action of G on $V^{\mathbb{C}}$ and the algebra $\mathbb{C}[V^{\mathbb{C}}]^G$ of G-invariant polynomials on $V^{\mathbb{C}}$. The generators $\sigma_1, \ldots, \sigma_n$ of the algebra $\mathbb{R}[V]^G$ are the generators of the algebra $\mathbb{C}[V^{\mathbb{C}}]^G$ as well. Let $\sigma^{\mathbb{C}} := (\sigma_1, \ldots, \sigma_n) : V^{\mathbb{C}} \to \mathbb{C}^m$ be the complexification of the map σ. Then $\sigma^{\mathbb{C}}(V^{\mathbb{C}})$ is canonically identified with the orbit space $V^{\mathbb{C}}/G$.

Note that if $G = W$ is a finite reflection group in V, the algebra $\mathbb{R}[V]^W$ is polynomial, i.e., the homogeneous generators $\sigma_1, \ldots, \sigma_n$ are algebraically independent, where $n = \dim V$, and, therefore, $\sigma^{\mathbb{C}}(V^{\mathbb{C}}) = \mathbb{C}^n$. The following theorem is due to Lyashko [**Lya83**].

Theorem. *Let W be a finite reflection group on V. Any biholomorphic diffeomorphism of the orbit space $V^{\mathbb{C}}/G = \mathbb{C}^n$ has a biholomorphic lift to $V^{\mathbb{C}}$.*

Lyashko considers only the irreducible Weyl groups and the germs of biholomorphic diffeomorphisms at 0. But his proof is evidently true for arbitrary finite reflection groups and for global biholomorphic diffeomorphisms.

16.1.4. Denote by W the subgroup of G generated by all reflections belonging to G. Since the projection $\sigma^{\mathbb{C}} : V^{\mathbb{C}} \to V^{\mathbb{C}}/G$ factors through the map $\sigma' : V^{\mathbb{C}} \to V^{\mathbb{C}}/W$ we have the following sequence of maps:

$$V^{\mathbb{C}} \xrightarrow{\sigma'} V^{\mathbb{C}}/W \xrightarrow{l} V^{\mathbb{C}}/G.$$

Since W is a finite reflection group, $V^{\mathbb{C}}/W = \sigma'(V^{\mathbb{C}}) = \mathbb{C}^n$.

Denote by $V^{\mathbb{C}}_{\mathrm{reg}}$ the set of regular points of the G-module $V^{\mathbb{C}}$. Evidently the map l is an étale morphism on $\sigma'(V^{\mathbb{C}}_{\mathrm{reg}})$ and then a covering map. Denote by S_2 the union of all strata of complex codimension ≥ 2 of the orbit space $V^{\mathbb{C}}/W$ and put $V_2 = (V^{\mathbb{C}}/W) \backslash S_2$.

Lemma. *V_2 is a universal covering space over $l(V_2)$.*

PROOF. If the group W is trivial, then $V_2 = V^{\mathbb{C}}_{\text{reg}}$ and our statement is true.

Suppose the group W is nontrivial, $x \in V_2 \backslash (V^{\mathbb{C}}_{\text{reg}}/W)$, and $v \in (\sigma')^{-1}(x)$. Then v belongs to some reflection hyperplane $H \subseteq V^{\mathbb{C}}$ and does not belong to the intersection of any two distinct reflection hyperplanes. Then there are some affine coordinates $z_1, \dots, z_n$ in a neighborhood U of v such that in U the maps σ' and $\sigma^{\mathbb{C}}$ have the same following expression:

$$(z_1, z_2, \dots, z_n) \to (z_1^2, z_2, \dots, z_n).$$

Since we may choose $u_1 = z_1^2, u_2 = z_2, \dots, u_n = z_n$ as local coordinates in some neighborhoods of $\sigma'(v)$ and $\sigma^{\mathbb{C}}(v)$, the map l is a diffeomorphism in $\sigma'(U)$. Then the restriction of l to V_2 is an étale map and thus a covering map onto $l(V_2) \subseteq V^{\mathbb{C}}/G$. Evidently the space V_2 is simply connected. □

16.1.5. Equivariant diffeomorphisms. Let $\operatorname{Aut}(G)$ be the group of automorphisms of the group G and S the set of reflections belonging to G. Denote by $\operatorname{Aut}(G, S)$ the group of automorphisms of G preserving the set S.

DEFINITION. A diffeomorphism F of V is called *equivariant with respect to* $a \in \operatorname{Aut}(G, S)$ (or just *a-equivariant*) if, for any $g \in G$, $F \circ g = a(g) \circ F$. If a is the identity of $\operatorname{Aut}(G, S)$, the diffeomorphism F is called equivariant.

Denote by H_s the reflection hyperplane of a reflection s. If a is a permutation of G and the diffeomorphism F of V satisfies the condition $F \circ g = a(g) \circ F$ for each $g \in G$, then $a \in \operatorname{Aut}(G)$ and $F(H_s)$ is the set of fixed points of $a(s)$. Therefore, $a(s) \in S$ and $a \in \operatorname{Aut}(G, S)$. Clearly any smooth a-equivariant diffeomorphism of V induces a smooth diffeomorphism of V/G.

For each $g \in G$, the diffeomorphism $x \mapsto gx$ $(x \in V)$ of V is an analytic a_g-equivariant diffeomorphism of V, where a_g is the inner automorphism $h \mapsto ghg^{-1}$ and this diffeomorphism induces the identity diffeomorphism of V/G.

16.1.6. THEOREM. *Suppose G is finite. Each smooth (real analytic) diffeomorphism f of the orbit space V/G has a smooth (real analytic) lift to V. Each lift F of f is an a-equivariant diffeomorphism of V with respect to some $a \in \operatorname{Aut}(G, S)$. If F_1 and F_2 are two lifts of the same smooth diffeomorphism of V/G there is a unique $g \in G$ such that $F_2 = g \circ F_2$.*

PROOF. By corollary 16.1.2, it is sufficient to consider only quasi-linear diffeomorphisms. Let $f = (f_1, \dots, f_n)$ be a quasi-linear diffeomorphism of V/G. Then f defines a quasi-linear diffeomorphism $f^{\mathbb{C}}$ of $V^{\mathbb{C}}/G$. Since $f^{\mathbb{C}}$ preserves codimensions of strata the restriction f' of $f^{\mathbb{C}}$ to $\sigma^{\mathbb{C}}(V_2)$ is a diffeomorphism of $\sigma^{\mathbb{C}}(V_2)$. Lemma 16.1.4 implies that f' has a holomorphic lift to $\sigma'(V_2)$ and by Hartog's extension theorem this lift has a holomorphic extension $\tilde{f}$ to the whole of $V^{\mathbb{C}}/W$. This extension is a lift of $f^{\mathbb{C}}$ to $V^{\mathbb{C}}/W$. It is clear that $\tilde{f}$ is a holomorphic diffeomorphism of $V^{\mathbb{C}}/W = \mathbb{C}^n$ since it is invertible. By theorem 16.1.3, the diffeomorphism $\tilde{f}$ has a biholomorphic lift $F^{\mathbb{C}}$ to $V^{\mathbb{C}}$. Since $F^{\mathbb{C}}(V_{\text{reg}}) \subseteq V_{\text{reg}}$, $F^{\mathbb{C}}(V) \subseteq V$. This means that the restriction of $F^{\mathbb{C}}$ to V is a lift of f to V.

Let $v_0 \in V_{\text{reg}}$ and $g \in G$. Then $F(g(v_0)) = a(g)(F(v_0))$, where a is a uniquely defined permutation of G. Since the map σ is an étale map on V_{reg} and V_{reg} is dense in V, the equality $F(g(v)) = a(g)(F(v))$ is true on the component of each $v_0 \in V_{\text{reg}}$ in V_{reg} and therefore by continuity on the whole of V. Then by the above remark $a \in \operatorname{Aut}(G, S)$.

The last statement of the theorem is evident. □

REMARKS. (1) Let $\rho : G \to \mathrm{O}(V)$ be an orthogonal representation of a finite group G and $\rho^{\mathbb{C}}$ the induced representation in $V^{\mathbb{C}}$. The sheaf of complex analytic

G-invariant functions on $V^{\mathbb{C}}$ defines a complex analytic structure on the orbit space $V^{\mathbb{C}}/G$. Let f be a biholomorphic diffeomorphism of $V^{\mathbb{C}}/G$. Then the existence of a biholomorphic lift for f is proved by the arguments in the proof of theorem 16.1.6.

(2) Let f be a local smooth diffeomorphism of the orbit space V/G defined in a neighborhood of $0 \in V/G$ such that $f(0) = 0$. The results obtained so far have local versions. In particular, for a germ f_0 of a local diffeomorphism f of V/G at 0 there is a germ of a local smooth diffeomorphism of V which is projectable to f_0.

16.1.7. The component group $\pi_0(\mathrm{Diff}(V/G))$ of $\mathrm{Diff}(V/G)$. Let now G be a compact Lie group and $\rho : G \to \mathrm{O}(V)$ its orthogonal representation. Let $\mathrm{Diff}_G(V)$ be the group of smooth diffeomorphisms of V which are projectable to smooth diffeomorphisms of V/G, i.e., the normalizer of G in the group $\mathrm{Diff}(V)$ of smooth diffeomorphisms of V. We may consider $\mathrm{Diff}_G(V)$ as a topological group with respect to the C^∞ topology.

Denote by $\mathrm{Diff}(V/G)$ the group of smooth diffeomorphisms of V/G. It is a quotient space of a subspace of the space of smooth maps of $\mathbb{R}^n$ into itself preserving $\sigma(V) \subseteq \mathbb{R}^n$. Since this space of smooth maps is a topological space with respect to C^∞ topology, we may endow $\mathrm{Diff}(V/G)$ with the quotient topology and consider $\mathrm{Diff}(V/G)$ as a topological group. We have the natural homomorphism of topological groups $\mathrm{Diff}_G(V) \to \mathrm{Diff}(V/G)$.

Let $\mathrm{Diff}_0(V/G)$ be the identity component of $\mathrm{Diff}(V/G)$ and let $\pi_0 = \pi_0(\mathrm{Diff}(V/G)) = \mathrm{Diff}(V/G)/\mathrm{Diff}_0(V/G)$ be the component group of $\mathrm{Diff}(V/G)$.

Denote by $P(V/G)$ the subgroup of $\mathrm{Diff}(V/G)$ consisting of quasi-linear diffeomorphisms of V/G. It is easily seen that $P(V/G)$ is a finite dimensional Lie group. Denote by $P_0(V/G)$ the identity component of $P(V/G)$.

Consider the normalizer of G in $\mathrm{GL}(V)$, namely, $N(G) = \mathrm{GL}(V) \cap \mathrm{Diff}_G(V)$, i.e., the group of a-equivariant automorphisms of the vector space V. It is clear that $N(G)$ is a finite dimensional Lie group.

16.1.8. Theorem. *We have:*

(1) *Let f be a quasi-linear diffeomorphism of V/G which has a lift to V. Then there is a lift F of f belonging to $N(G)$.*
(2) $\pi_0 = P(V/G)/P_0(V/G)$.
(3) *For a finite group G, the group π_0 is isomorphic to the component group of the group $N(G)/G$.*

Proof. To (1): For each $v \in V$, we get $\sigma(\frac{1}{t}F(tv)) = \sigma(F(v))$. Then $F_0 = \lim_{t\to 0} \frac{1}{t}F(tv)$, i.e., the derivative of F at 0, is a lift of f belonging to $N(G)$.

To (2): By the proof of theorem 16.1.2, the groups $P(V/G)$ and $\mathrm{Diff}(V/G)$ are homotopically equivalent and then $\pi_0 = P(V/G)/P_0(V/G)$.

To (3): By theorem 16.1.6 and (1), for a finite group G, we have $P(V/G) = N(G)/G$. □

16.1.9. Denote by $Z(G)$ the centralizer of G in $\mathrm{GL}(V)$ and by $Z'(G)$ the quotient group $Z(G)/(Z(G) \cap G)$. Put $\pi_0' = Z'(G)/Z_0'(G)$, where $Z_0'(G)$ is the component of the identity of $Z'(G)$.

Theorem. *For a finite group G we have:*

(1) *The quotient group $(N(G)/G)/Z'(G)$ is isomorphic to a subgroup of the group $\mathrm{Aut}(G,S)/\mathrm{Int}(G)$. In other words, the group π_0 is an extension of a subgroup of $\mathrm{Aut}(G,S)/\mathrm{Int}(G)$ with the kernel isomorphic to π_0'.*
(2) *$Z'(G)$ is an open normal subgroup of $N(G)/G$.*
(3) *If the G-module V is irreducible, the group π_0' is either trivial or isomorphic to $\mathbb{Z}_2$. Then the group π_0 is finite.*

PROOF. By 16.1.5, there is a natural map $N(G) \to \mathrm{Aut}(G,S)$. It is easily checked that this map is a group homomorphism, its kernel equals $Z(G)$ and the image of G in $N(G)$ equals $\mathrm{Int}(G)$. Then the quotient group $(N(G)/G)/Z'(G)$ is isomorphic to a subgroup of the group $\mathrm{Aut}(G,S)/\,\mathrm{Int}(G)$.

Since the group G is finite for each $a \in \mathrm{Aut}(G,S)$ the set N_a of a-equivariant diffeomorphisms from $N(G)$ is an open subset of $N(G)$ and $N(G)$ is a disjoint union of the non-empty N_a. In particular, $Z(G)$ is an open normal subgroup of $N(G)$. This implies that $Z'(G)$ is an open normal subgroup of $N(G)/G$.

To prove (3) note that there are the following three possibilities: Either $Z(G) = \mathbb{R}\backslash\{0\}$, or $Z(G) = \mathbb{C}\backslash\{0\}$, or $Z(G) = \mathbb{H}\backslash\{0\}$, where $\mathbb{H}$ is the algebra of quaternions. $\square$

Theses results are evidently also true for the group of real analytic diffeomorphisms of the orbit space V/G.

16.1.10. In [**Los01**] the groups $\pi_0(\mathrm{Diff}(V/W))$ and $\pi_0(\mathrm{Diff}(V^p/W))$ for a finite reflection group W and the diagonal action of W in V^p are calculated by means of the automorphism group of the Coxeter diagram of W.

16.2. Lifting mappings over invariants

We consider a representation of a finite group G in a complex vector space V. In this case the orbit space $Z = V/G$ coincides with the categorical quotient $V/\!\!/G$ which is a normal affine variety. Therefore the orbit space Z has the natural structure of a complex analytic set and there are several types of morphisms into $V/\!\!/G$, like regular, rational, or holomorphic. The conditions of lifting for holomorphic automorphisms of orbit spaces were found for the Weyl groups in [**Lya83**] and for arbitrary finite groups in [**KLM03**]. In [**LMP03**] it was proved that each holomorphic lift of a regular automorphism of the orbit space is regular.

In [**KLMR08**] the conditions for lifts of germs of holomorphic morphisms at 0 from $\mathbb{C}^p$ to Z, for lifts of regular maps from $\mathbb{C}^p$ to Z, and for lifts of formal morphisms from $\mathbb{C}^p$ to Z, i.e., the morphisms of the $\mathbb{C}$-algebra $\mathbb{C}[Z]$ to the ring of formal power series $\mathbb{C}[[X_1,\dots,X_p]]$ in variables $X_1,\dots,X_p$, were considered. We shall present those results.

16.2.1. Preliminaries. Let V be an n-dimensional complex vector space, G a finite subgroup of $\mathrm{GL}(V)$, and $\mathbb{C}[V]^G$ the algebra of G-invariant polynomials on V.

The following facts are well known (see, e.g. [**VP89**]). Denote by Z the categorical quotient $V/\!\!/G$, i.e., the normal affine algebraic variety with the coordinate ring $\mathbb{C}[V]^G$. Since the group G is finite, the categorical quotient $V/\!\!/G$ is the geometric one, i.e., $V/\!\!/G$ is the orbit space V/G. Let $\pi = \pi_V : V \to Z$ be the quotient projection. The affine algebraic variety Z has the natural structure of a complex analytic space: Let $\sigma_1,\dots,\sigma_m$ be a minimal system of homogeneous generators of the algebra $\mathbb{C}[V]^G$ and let $\sigma : V \to \sigma(V) \subseteq \mathbb{C}^m$ be the corresponding morphism. Then $\sigma(V)$ is an irreducible Zariski-closed subset of $\mathbb{C}^n$ which is isomorphic to the affine variety Z. For this presentation of Z the morphism $\sigma : V \to \sigma(V)$ coincides with the projection π.

In the sequel we assume that a minimal system of homogeneous generators $\sigma_1,\dots,\sigma_m$ is fixed.

Let K be a subgroup of G. We denote by $V_{(K)}$ the set of points of V whose isotropy groups are conjugate to K. By definition, $V_{(K)} \subseteq \cup_{g\in G} V^{gKg^{-1}}$, where V^K is the subspace of V of fixed points of the action of K on V. Put $Z_{(K)} := \pi(V_{(K)})$. It is known that $\{Z_{(K)} : K < G\}$ is a finite stratification of Z into locally closed irreducible smooth algebraic subvarieties. This is the simplest case of a Luna stratification, see [**Lun73**]. Put $V_0 := V_{(K)}$ for $K = \{\mathrm{id}\}$ and $Z_0 := \pi(V_0)$.

Denote by $Z_{>i}$ the union of the strata of codimension greater than i and put $Z_{\leq i} := Z \backslash Z_{>i}$. Then $Z_{>i}$ is a Zariski-closed subset of Z and $Z_0 = Z_{\leq 0} = Z \backslash Z_{>0}$ is a stratum of Z called the principal stratum. Points in Z_0 and in V_0 are called regular points. The following proposition is evident.

PROPOSITION. *Z_0 is a Zariski-open smooth subvariety of Z and the restriction of π to V_0 is an étale morphism onto Z_0.* □

16.2.2. Invariant coordinates. For each regular point $z_0 \in Z_0$ there is a system of regular functions $z_1, \dots, z_n$ on Z such that each $y_i := z_i \circ \pi$ equals one of the generators σ_j and the functions $z_i - z_i(z_0)$ are local parameters at z_0. Then the y_i are local coordinates on V in a neighborhood of each point $v \in \pi^{-1}(z_0)$. By definition, the functions y_i are G-invariant. These coordinates y_i are called *invariant coordinates* on V. Since we fixed the basic generators $\sigma_1, \dots, \sigma_m$, there are only finitely many choices of such invariant coordinates on V.

Let e_i, $1 \leq i \leq n$, be a basis of V and u_i the corresponding coordinates on V. Denote by $J \in \mathbb{C}[V]$ the Jacobian $\det(\frac{\partial y_i}{\partial u_j})$. It is clear that J is a homogeneous polynomial.

PROPOSITION. *For each integer $k > 0$ there is a G-invariant polynomial $\Delta_k \in \mathbb{C}[V]^G$ of minimal degree such that J^k divides Δ_k and the sets of zeros of J and Δ_k coincide. The polynomial Δ_k is unique up to a non-zero factor $c \in \mathbb{C}$.*

PROOF. Let $J = f_1^{n_1} \dots f_s^{n_s}$ be a decomposition of J into the product of linearly independent irreducible polynomials $f_l \in \mathbb{C}[V]$. Consider the principal effective divisor $(J) = n_1(f_1) + \dots + n_s(f_s)$ of the polynomial J on V. Since for each $g \in G$ we have $J \circ g = \det(g_j^i).J$, where (g_j^i) is the matrix of g in the basis e_i, the divisor (J) is G-invariant. Then each $g \in G$ permutes the prime divisors (f_p), $1 \leq p \leq s$. This implies that, if $g(f_p) = (f_q)$, the coefficients n_p and n_q of the divisor (J) are equal. Let $\{m_1, \dots, m_l\}$ be the set of distinct coefficients of the divisor (J) and let, for each m_α, Φ_α be the product of distinct factors f_p of J having the same power m_α in the above decomposition of J. Then we have $J = \prod_{\alpha=1}^l \Phi_\alpha^{m_\alpha}$. By the above arguments, for each $1 \leq \alpha \leq l$, the divisor (Φ_α) of the polynomial Φ_α is G-invariant.

Since the group G is finite, for each $1 \leq \alpha \leq l$, there is a minimal integer $p_\alpha > 0$ such that the polynomial $\Phi_\alpha^{p_\alpha}$ is G-invariant. For $1 \leq \alpha \leq l$, let $km_\alpha = s_\alpha p_\alpha + r_\alpha$, where s_α and r_α are unique non-negative integers such that $0 \leq r_\alpha < p_\alpha$. Then we have $J^k = \prod_{\alpha=1}^l \Phi_\alpha^{s_\alpha p_\alpha + r_\alpha}$. Let μ_α be the least common multiple of r_α and p_α. Then $\Delta_k = \prod_{\alpha=1}^l \Phi_\alpha^{s_\alpha p_\alpha + \mu_\alpha}$ is a G-invariant polynomial of minimal degree such that the sets of zeros of J and Δ_k coincide and J^k divides Δ_k.

By the above formula for J^k, for each G-invariant polynomial P such that the sets of zeros of J and P coincide and J^k divides P, Δ_k divides P. □

We denote by $\tilde{\Delta}_k$ the regular function on Z such that $\tilde{\Delta}_k \circ \pi = \Delta_k$. By definition, we have $\tilde{\Delta}_k(z_0) \neq 0$ for each k and $z_0 \in Z_0$. Conversely, let y_i be invariant coordinates on V, let z_i be the corresponding regular functions on Z, and for some positive integer k let $\tilde{\Delta}_k$ be the corresponding regular function on Z. If, for a point $z \in Z$, we have $\tilde{\Delta}_k(z) \neq 0$, then $z \in Z_0$.

Later, for the sake of simplicity, we put $\Delta := \Delta_1$ and $\tilde{\Delta} := \tilde{\Delta}_1$.

Denote by $V(\tilde{\Delta})$ the set of zeros of $\tilde{\Delta}$. Thus $Z_{>0}$ is the intersection of the Zariski-closed subsets $V(\tilde{\Delta})$ obtained from all choices of invariant coordinates constructed from the basic generators of $\mathbb{C}[V]^G$. The similar statement is true for $V \backslash V_0$ if we replace $\tilde{\Delta}$ by Δ.

16.2.3. Jet spaces. For an affine variety X, we will define the space $J_0^q(\mathbb{C}^p, X)$ of q-jets at 0 of morphisms from $\mathbb{C}^p$ to X.

Consider the category of $\mathbb{C}$-algebras. Let $\mathbb{C}[X_1, \ldots, X_p]$ be the $\mathbb{C}$-algebra of polynomials in variables $X_1, \ldots, X_p$ with complex coefficients and let $\mathfrak{m}_p$ be the ideal in $\mathbb{C}[X_1, \ldots, X_p]$ generated by the $X_1, \ldots, X_p$. Put $\mathfrak{m}_p^q := (\mathfrak{m}_p)^q$. Then

$$\mathfrak{J}_p^q := \mathbb{C}[X_1, \ldots, X_p]/\mathfrak{m}_p^{q+1}$$

is the truncated ring of polynomials, the model jet algebra. In particular, $\mathfrak{J}_p^0 = \mathbb{C}[X_1, \ldots, X_p]/\mathfrak{m}_p = \mathbb{C}$.

Let $A = (a_1, \ldots, a_s)$ for $a_1, \ldots, a_s \in \{1, \ldots, p\}$ be a (unordered) multi-index of order $|A| := s$. In particular, for $s = 0$ we put $A := \emptyset$. Denote by $\mathfrak{A}_{p,q}$ the set of multi-indices A of orders $\leq q$. By definition, each $P \in \mathfrak{J}_p^q$ can be written as $P = \sum_{A \in \mathfrak{A}_{p,q}} p_A\, X_A$, where $p_A \in \mathbb{C}$ and $X_A := X_{a_1} \ldots X_{a_s}$. The natural bijection $P \mapsto (p_A)_{A \in \mathfrak{A}_{p,q}}$ between $\mathfrak{J}_p^q$ and $\mathbb{C}^{\mathfrak{A}_{p,q}}$ is an isomorphism of vector spaces and defines a structure of affine space on $\mathfrak{J}_p^q$. For $q \leq r$, consider the natural morphism $\rho_{r,q} : \mathfrak{J}_p^r \to \mathfrak{J}_p^q$.

For an affine variety X over $\mathbb{C}$, the set of $\mathfrak{J}_p^q$-valued points of X, i.e., morphisms from the coordinate ring $\mathbb{C}[X]$ of X to the ring $\mathfrak{J}_p^q$, is called the *space of q-jets* of morphisms from $\mathbb{C}^p$ to X at $0 \in \mathbb{C}^p$ and is denoted by $J_0^q(\mathbb{C}^p, X)$. In particular, we have $J_0^0(\mathbb{C}^p, X) = X$ and $J_0^1(\mathbb{C}^p, X) = TX$, the total tangent bundle of X.

It is evident that each polynomial function on $\mathfrak{J}_p^q$ defines a function on $J_0^q(\mathbb{C}^p, X)$ and these functions generate a ring of $\mathbb{C}$-valued functions on $J_0^q(\mathbb{C}^p, X)$. It is clear that this ring is a finitely generated $\mathbb{C}$-algebra. Then $J_0^q(\mathbb{C}^p, X)$ supplied with this ring has a structure of an affine variety (not necessarily irreducible). For two affine varieties X_1 and X_2 and for a morphism $\varphi : X_1 \to X_2$, we have the natural morphism $J_0^q(\mathbb{C}^p, \varphi) : \bar{J}_0^q(\mathbb{C}^p, X_1) \to \bar{J}_0^q(\mathbb{C}^p, X_2)$ of affine varieties. Thus one can consider $J_0^q(\mathbb{C}^p, \quad)$ as a covariant functor from the category of affine varieties to itself.

For each $h \in J_0^q(\mathbb{C}^p, X)$ there is a unique point $x \in X$ such that the corresponding maximal ideal $\mathfrak{m}_x$ coincides with the kernel of the composition $\rho_{q,0} \circ h$. Then the morphism h can be extended uniquely to a morphism from $\mathcal{O}_x$, the local ring of x, to $\mathfrak{J}_p^q$ vanishing on $\mathfrak{m}_x^{q+1}$ and hence induces a morphism $h_{x,q} : \mathcal{O}_x/\mathfrak{m}_x^{q+1} \to \mathfrak{J}_p^q$ which, in turn, determines the initial morphism h uniquely. Therefore one can view $J_0^q(\mathbb{C}^p, X)$ as the set of morphisms from the local rings $\mathcal{O}_x/\mathfrak{m}_x^{q+1}$ to $\mathfrak{J}_p^q$ for all $x \in X$.

$$\begin{tikzcd}
& \mathbb{C}[X] \arrow[r, "h"] \arrow[d] & \mathfrak{J}_p^q \\
\mathfrak{m}_x^{q+1} \arrow[r, hook] & \mathcal{O}_x \arrow[r] \arrow[ur, dashed, "h"] & \mathcal{O}_x/\mathfrak{m}_x^{q+1} \arrow[u, dashed, "h_{x,q}"']
\end{tikzcd}$$

Assume that X is presented as a Zariski-closed subset of $\mathbb{C}^m$ defined by an ideal $(\Phi_1, \ldots, \Phi_r)$ of the ring $\mathbb{C}[W_1, \ldots, W_m]$ of polynomials with complex coefficients in variables $W_1, \ldots, W_m$.

A morphism $h : \mathbb{C}[X] \to \mathfrak{J}_p^q$ is defined by a morphism $h' : \mathbb{C}[W_1, \ldots, W_m] \to \mathfrak{J}_p^q$ with $h'(\Phi_l) = 0$ for $1 \leq l \leq r$. It is determined by $h'(W_i) = \sum_{A \in \mathfrak{A}_{p,q}} W_{i,A}\, X_A$, where $1 \leq i \leq m$ and $W_{i,A} \in \mathbb{C}$. The condition $h'(\Phi_l) = 0$ is equivalent to the vanishing of all the coefficients of the variables X_A in $h(\Phi_l)$. Thus the map $J_0^q(\mathbb{C}^p, X) \to (\mathbb{C}^m)^{\mathfrak{A}_{p,q}}$ given by $h \mapsto (W_{i,A})_{1 \leq i \leq m,\, A \in \mathfrak{A}_{p,q}}$ induces a bijective correspondence between $J_0^q(\mathbb{C}^p, X)$ and the Zariski-closed subset of $(\mathbb{C}^m)^{\mathfrak{A}_{p,q}}$ defined by $r|\mathfrak{A}_{p,q}|$ many polynomial equations, where $|\mathfrak{A}_{p,q}|$ denotes the cardinality of the set $\mathfrak{A}_{p,q}$. By definition, this correspondence is an isomorphism of affine varieties.

The homomorphism $\rho_{r,q}$ induces the morphism

$$p_{X,r,q} : J_0^r(\mathbb{C}^p, X) \to J_0^q(\mathbb{C}^p, X).$$

In particular, we have the morphism $p_{X,q,0} : J_0^q(\mathbb{C}^p, X) \to X$.

The projective limit $J_0^\infty(\mathbb{C}^p, X) = \varprojlim J_0^q(\mathbb{C}^p, X)$ is called the *space of* ∞-*jets* at $0 \in \mathbb{C}^p$ of morphisms from $\mathbb{C}^p$ to X or the *space of formal morphisms* from $\mathbb{C}^p$ to X. By the definition of a projective limit we have natural projections $p_{X,\infty,q} : J_0^\infty(\mathbb{C}^p, X) \to J_0^q(\mathbb{C}^p, X)$. By definition, one can consider a point of $J_0^\infty(\mathbb{C}^p, X)$ either as a morphism $\mathbb{C}[X] \to \mathbb{C}[[X_1, \dots, X_p]]$ or as a morphism from the completion $\tilde{\mathcal{O}}_x$ of the local ring $\mathcal{O}_x$ for some $x \in X$ to $\mathbb{C}[[X_1, \dots, X_p]]$.

In particular, for the above presentation of X as a Zariski-closed subset of $\mathbb{C}^m$, each $h \in J_0^\infty(\mathbb{C}^p, X)$ is uniquely defined by a morphism $h' : \mathbb{C}[W_1, \dots, W_m] \to \mathbb{C}[[X_1, \dots, X_p]]$ with $h'(\Phi_l) = 0$ for $1 \le l \le r$. It is defined by $h'(W_i) = h_i \in \mathbb{C}[[X_1, \dots, X_p]]$, where $\Phi_l(h_i) = 0$ for each $1 \le l \le r$.

Consider X as a complex analytic set and let $h : \mathbb{C}^p, 0 \to X$ be a germ of a holomorphic map at $0 \in \mathbb{C}^p$. Denote by $\mathfrak{F}_{\mathbb{C}^p,0}$ and $\mathfrak{F}_{X,x}$ the rings of germs of holomorphic functions on $\mathbb{C}^p$ at 0 and on X at x respectively. We may identify the ring $\mathfrak{F}_{\mathbb{C}^p,0}$ with a subring of the ring $\mathbb{C}[[X_1, \dots, X_p]]$. Consider the morphism $h^* : \mathfrak{F}_{X,x} \to \mathfrak{F}_{\mathbb{C}^p,0} \subseteq \mathbb{C}[[X_1, \dots, X_p]]$ induced by h. The restriction of h^* to $\mathcal{O}_{X,x}$, which is denoted by $j_0^\infty h$, belongs to $J_0^\infty(\mathbb{C}^p, X)$ and is called the ∞-jet of h at 0. Put $j_0^q h := p_{X,\infty,q}(j_0^\infty h)$ and call $j_0^q h$ the q-jet of h at 0.

Denote by $x_1, \dots, x_p$ the standard coordinates in $\mathbb{C}^p$. Let $A = (a_1, \dots, a_s)$ be a multi-index, W a finite dimensional complex vector space, and $F : \mathbb{C}^p, 0 \to W$ a germ of a holomorphic map, i.e., $F \in W \otimes \mathfrak{F}_{\mathbb{C}^p,0}$. We denote by ∂_A a linear operator on $W \otimes \mathfrak{F}_{\mathbb{C}^p,0}$ which is equal to the tensor product of the identical operator on W and the operator $\frac{\partial^{|A|}}{\partial x_{a_1} \dots \partial x_{a_s}}$ on $\mathfrak{F}_{\mathbb{C}^p,0}$. In particular, we have $\partial_\emptyset F = F$ and we write $\partial_a F$ instead of $\partial_{(a)} F$.

For the above presentation of X as a Zariski-closed subset in $\mathbb{C}^m$ the holomorphic germ $h : \mathbb{C}^p, 0 \to X$ can be given by a holomorphic map F from a neighborhood of $0 \in \mathbb{C}^p$ to $\mathbb{C}^m$ such that $\Phi_l \circ F = 0$ for each $1 \le l \le r$. Denote by $\mathfrak{A}_p$ the set of all multi-indices $A = (a_1, \dots, a_s)$. By definition, the ∞-jet $j_0^\infty h$ is uniquely determined by the indexed set $(\partial_A F(0))_{A \in \mathfrak{A}_p}$ of complex numbers. The complex numbers $\partial_A F(0)$ satisfy the equations $\partial_A(\Phi_l \circ F)(0) = 0$ for $A \in \mathfrak{A}_p$ and do not depend on the choice of F. Similarly, the q-jet $j_0^q F$ is determined by the indexed set $(\partial_A F(0))_{A \in \mathfrak{A}_{p,q}}$ of complex numbers satisfying the equations $\partial_A(\Phi_l \circ F)(0) = 0$ for all $A \in \mathfrak{A}_{p,q}$. The above considerations show that for a smooth point $x \in X$ our notion of jets coincide with the usual one.

Note that the jet spaces of holomorphic functions and of regular functions on affine varieties coincide.

Later we denote by ∂_A also the linear operator on $W \otimes \mathbb{C}[[X_1, \dots, X_p]]$ which is equal to the tensor product of the identical operator on W and the operator $\frac{\partial^{|A|}}{\partial x_{a_1} \dots \partial x_{a_s}}$ on $\mathbb{C}[[X_1, \dots, X_p]]$.

16.2.4. Consider a G-module V, the spaces of q-jets $J_0^q(\mathbb{C}^p, V)$, and $J_0^q(\mathbb{C}^p, Z)$, and the sets of formal morphisms $J_0^\infty(\mathbb{C}^p, V)$ and $J_0^\infty(\mathbb{C}^p, Z)$. The projection $\pi : V \to Z$ induces the morphism $J_0^q(\mathbb{C}^p, \pi) : J_0^q(\mathbb{C}^p, V) \to J_0^q(\mathbb{C}^p, Z)$ and the map $J_0^\infty(\mathbb{C}^p, \pi) : J_0^\infty(\mathbb{C}^p, V) \to J_0^\infty(\mathbb{C}^p, Z)$.

The standard action of the group G on $\mathbb{C}[V]$ induces an action of G on $J_0^q(\mathbb{C}^p, V)$ by automorphisms of $J_0^q(\mathbb{C}^p, V)$ as an affine variety and on $J_0^\infty(\mathbb{C}^p, V)$. The inclusion $\mathbb{C}[V]^G \subseteq \mathbb{C}[V]$, the morphism $J_0^q(\mathbb{C}^p, \pi)$, and the map $J_0^\infty(\mathbb{C}^p, \pi)$ induce the morphism

$$\pi^q : J_0^q(\mathbb{C}^p, V)/G \to J_0^q(\mathbb{C}^p, Z)$$

and the map

$$\pi^\infty : J_0^\infty(\mathbb{C}^p, V)/G \to J_0^\infty(\mathbb{C}^p, Z).$$

Denote by $\bar{J}_0^q(\mathbb{C}^p, Z)$ the Zariski closure of $\pi^q(J_0^q(\mathbb{C}^p, V)/G)$ in $J_0^q(\mathbb{C}^p, Z)$.

PROPOSITION. *The morphism $\pi^q : J_0^q(\mathbb{C}^p, V)/G \to J_0^q(\mathbb{C}^p, Z)$ induces a birational morphism of $J_0^q(\mathbb{C}^p, V)/G$ to $\bar{J}_0^q(\mathbb{C}^p, Z)$.*

PROOF. The group G acts freely on the open subset $p_{V,q,0}^{-1}(V_0) \subseteq J_0^q(\mathbb{C}^p, V)$. Since all points of V_0 and Z_0 are smooth, for each $v \in V_0$ the morphism $J_0^q(\mathbb{C}^p, \pi)$ induces a bijective map of $p_{V,q,0}^{-1}(v)$ onto $p_{Z,q,0}^{-1}(\pi(v))$. Thus the morphism π^q maps the Zariski-open subset $p_{V,q,0}^{-1}(V_0)/G$ of $J_0^q(\mathbb{C}^p, V)/G$ onto the Zariski-open subset $p_{Z,q,0}^{-1}(Z_0)$ of $J_0^q(\mathbb{C}^p, Z)$ bijectively and this implies the statement of the proposition. □

16.2.5. Evidently we have the following bijections

$$J_0^q(\mathbb{C}^p, V) = \mathrm{Hom}(\mathbb{C}[V], \mathfrak{J}_p^q) = \mathrm{Lin}(V^*, \mathfrak{J}_p^q) = V \otimes \mathfrak{J}_p^q,$$

where V^* is the dual vector space for V, 'Hom' means the set of morphisms in the category of $\mathbb{C}$-algebras, and 'Lin' means the set of linear mappings. So each $h \in J_0^q(\mathbb{C}^p, V) = V \otimes (\mathfrak{J}_p^q)$ can be written uniquely as $h = \sum_{A \in \mathfrak{A}_{p,q}} h_A \otimes X_A$, where $h_A \in V$. Similarly, $J_0^\infty(\mathbb{C}^p, V) = V \otimes \mathbb{C}[[X_1, \dots, X_p]]$ and, for $h \in J_0^\infty(\mathbb{C}^p, V)$, we have $h = \sum_{A \in \mathfrak{A}_p} h_A \otimes X_A$, where $h_A \in V$.

PROPOSITION. *The space of jets $J_0^q(\mathbb{C}^p, V)$ is isomorphic to the affine space $V^{\mathfrak{A}_{p,q}}$. Moreover, there are isomorphisms of G-modules*

$$J_0^q(\mathbb{C}^p, V) = V^{\mathfrak{A}_{p,q}}, \quad \textit{and} \quad J_0^\infty(\mathbb{C}^p, V) = V^{\mathfrak{A}_p}$$

where the G-action on the products is the diagonal one.

PROOF. The first statement follows from the definition of the structure of the affine variety on $J_0^q(\mathbb{C}^p, V)$. The maps $J_0^q(\mathbb{C}^p, V) \ni h \mapsto (h_A)_{A \in \mathfrak{A}_{p,q}}$ and $J_0^\infty(\mathbb{C}^p, V) \ni h \mapsto (h_A)_{A \in \mathfrak{A}_p}$ give the required isomorphisms of G-modules. □

Note that for

$$A = (\underbrace{1, \dots, 1}_{r_1 \text{ times}}, \dots, \underbrace{p, \dots, p}_{r_p \text{ times}}),$$

where $r_1, \dots, r_p \geq 0$, and for a germ of a holomorphic map $F : \mathbb{C}^p, 0 \to V$ we have $h = j_0^\infty F = \sum_A \frac{1}{A!} \partial_A F(0) \otimes X_A$, where $A! = r_1! \dots r_p!$. A similar formula is true for the q-jet $j_0^q F$.

16.2.6. The lifting problem. We consider the following problem. Let $f : \mathbb{C}^p \to Z$ be a rational morphism which is regular on a Zariski-open subset U of $\mathbb{C}^p$. A rational morphism $F : \mathbb{C}^p \to V$ which is regular on U is called a *regular lift* of f if $\pi \circ F = f$.

$$\begin{tikzcd} & V \arrow[d, "\pi"] \\ \mathbb{C}^p \arrow[ur, "F"] \arrow[r, "f"] & Z \end{tikzcd}$$

Similarly, let U be a connected classically open subset U of $\mathbb{C}^p$ and let $U \to Z$ be a holomorphic map, i.e., a morphism in the category of complex analytic sets. A holomorphic map $F : U \to V$ is called a *holomorphic lift* of f if $\pi \circ F = f$. If $f : \mathbb{C}^p, x \to Z$ is a germ at $x \in \mathbb{C}^p$ of a holomorphic map from $\mathbb{C}^p$ to Z, a germ $F : \mathbb{C}^p, x \to V$ at x of a holomorphic map from $\mathbb{C}^p$ to V is called a *local lift* of f if $\pi \circ F = f$.

Let $f \in J_0^\infty(\mathbb{C}^p, Z)$ be a formal morphism. A formal morphism $F \in J_0^\infty(\mathbb{C}^p, V)$ is called a *formal lift* of f if $J_0^\infty(\mathbb{C}^p, \pi) \circ F = f$.

A germ $f : \mathbb{C}^p, x \to Z$ at $x \in \mathbb{C}^p$ of a holomorphic map from $\mathbb{C}^p$ to Z is called *quasi-regular* if $f^{-1}(Z_0)$ meets any neighborhood of x.

By 16.2.2, if a germ f is quasi-regular, there is a choice of invariant coordinates y_i such that for the corresponding $\tilde{\Delta}$ the composition $\tilde{\Delta} \circ f$ does not vanish identically.

A formal morphism $f : \mathbb{C}[Z] \to \mathbb{C}[[X_1, \dots, X_p]]$ from $\mathbb{C}^p$ to Z is called *quasi-regular* if there is no stratum S of Z of codimension ≥ 1 such that f factors through a morphism $\mathbb{C}[\overline{S}] \to \mathbb{C}[[X_1, \dots, X_p]]$, where $\overline{S}$ is the closure of S.

We claim that *for a quasi-regular formal morphism f and for any choice of invariant coordinates y_i, for the corresponding $\tilde{\Delta}$ we have* $f(\tilde{\Delta}) \neq 0$. By 16.2.2, the set of zeros of all $\tilde{\Delta}$ obtained from any choice of invariant coordinates coincides with the Zariski-closed subset $Z_{>0}$ of Z. If our claim is wrong, by Hilbert's Nullstellensatz the formal morphism f vanishes on the ideal $I(Z_{>0}) = \sqrt{I(Z_{>0})}$ of $\mathbb{C}[W_1, \dots, W_m]$ which defines $Z_{>0}$. Consider the standard presentation of $I(Z_{>0})$ as an intersection of a finite set of prime ideals of $\mathbb{C}[Z]$ corresponding to the decomposition of $Z_{>0}$ into irreducible components. Since $\mathbb{C}[[X_1, \dots, X_p]]$ is an integral domain, f vanishes on at least one of these prime ideals. But each of these prime ideals defines a component of $Z_{>0}$ and such components are the closures of strata of Z of codimension ≥ 1. This contradicts our assumption.

Let K be a subgroup of G and let V^K be a subspace of V of fixed points of the action of K on V. Let $H = N_G(K)$ be the normalizer of K in G consisting of all elements of G preserving V^K, and let $W = H/K$ be the corresponding quotient group acting naturally on V^K. By definition, $\pi(V^K)$ is the closure $\overline{S}$ of a stratum S of Z.

Consider the natural map $\kappa : V^K/W \to \pi(V^K) = \overline{S}$. It is evidently bijective and regular since the natural map $V/W \to V/G$ is regular and $\pi(V^K) = \overline{S}$ is a Zariski-closed subset of Z as the projection π is a finite morphism. Then the morphism κ is birational. Denote by $\overline{S}_{\text{nor}}$ the set of all normal points of $\overline{S}$, i.e., points $x \in \overline{S}$ such that the local ring $\mathcal{O}_x(\overline{S})$ is integrally closed. It is known that $\overline{S}_{\text{nor}}$ is a Zariski-open subset of $\overline{S}$ and $S \subseteq \overline{S}_{\text{nor}}$ since S is smooth. Since the affine variety V^K/W is normal, by Zariski's main theorem, the restriction of κ to $\overline{S}_{\text{nor}}$ induces an isomorphism between the algebraic varieties $\pi^{-1}(\overline{S}_{\text{nor}})$ and $\overline{S}_{\text{nor}}$.

Assume that, for a holomorphic map f as above, $f(U)$ is contained in $Z_{>0}$. Then $f(U)$ is contained in the closure $\overline{S}$ of a stratum S of Z of codimension ≥ 1. Namely, let $f(U) \subseteq Z_{>i-1}$ for maximal i. Then there exists $x \in U$ such that $f(x)$ is a point of some stratum S of codimension i; otherwise $f(U) \subseteq Z_{>i}$. If a regular function $h \in \mathbb{C}[Z]$ vanishes on S then $h \circ f$ vanishes on an open neighborhood of x in U and thus on the whole of U. So $f(U) \subseteq \overline{S}$ and there is a subgroup K of G distinct from G such that $f(U) \subseteq \overline{S} = \pi(V^K)$.

It is clear that if each morphism f of the above type (regular, holomorphic, or formal) from $\mathbb{C}^p$ to V^K/W has a lift F (regular, holomorphic, local, or formal), then the composition of f with the morphism κ has the corresponding lift to V which is the composition of F with the inclusion $V^K \to V$.

Conversely, if $f : \mathbb{C}, x \to Z$ is a germ of a holomorphic map at $x \in \mathbb{C}^p$ such that $f(x) \in \overline{S}_{\text{nor}}$, there is a unique germ of a holomorphic map $f' : \mathbb{C}^p, x \to V^K/W$ such that $\kappa \circ f' = f$. Similarly, let $f : \mathbb{C}[Z] \to \mathbb{C}[[X_1, \dots, X_p]]$ be a formal morphism from $\mathbb{C}^p$ to Z which can be extended to the morphism $\tilde{\mathcal{O}}_z(Z) \to \mathbb{C}[[X_1, \dots, X_p]]$ for some $z \in \overline{S}_{\text{nor}}$. There is a unique formal morphism $f' : \mathbb{C}[V/W] \to \mathbb{C}[[X_1, \dots, X_p]]$ such that $J_0^\infty(\mathbb{C}^p, \kappa)(f') = f$. In both cases the lifting problem for $\pi : V \to V/G$ reduces to the corresponding one for $\pi_{V^K} : V^K \to V^K/W$.

Although these arguments give nothing if the above points $f(x)$ or z do not belong to $\overline{S}_{\mathrm{nor}}$, we have the following theorem.

THEOREM. *We have:*

(1) *Let f be a holomorphic map from a classically open subset U of $\mathbb{C}^p$ to $Z = V/G$, S a stratum of maximal codimension such that $f(U) \subseteq \overline{S}$, and $f : \mathbb{C}^p, x \to Z$ be a germ of f at some $x \in U$. Let K be a subgroup of G such that $\pi(V^K) = \overline{S}$, and let $W = N_G(K)/K$. If the germ $f : \mathbb{C}^p, x \to Z$ at $x \in U$ has a local lift to V, then the germ $f' : \mathbb{C}^p, x \to V^K/W$ of the map defined by f is a quasi-regular germ of a holomorphic map and this germ has a local lift to V^K.*

$$\begin{tikzcd}
& & V^K \arrow[r, hook] \arrow[dl, two heads] \arrow[dd, two heads, "\pi"] & V \arrow[dd, two heads, "\pi"] & \\
& V^K/W \arrow[dr, "\kappa"] & & & \\
\mathbb{C}^p, x \arrow[ur, "f'"] \arrow[rr, "f"'] & & \overline{S} \arrow[r, hook] & Z \arrow[r, equal] & V/G
\end{tikzcd}$$

(2) *Let f be a formal morphism from $\mathbb{C}^p$ to Z and let S be a stratum of maximal codimension such that f factors through a formal morphism f' from $\mathbb{C}^p$ to $\overline{S}$. If f has a formal lift to V, then the formal morphism f' is quasi-regular and has a formal lift to V^K such that $\pi(V^K) = \overline{S}$.*

(3) *If F_1 and F_2 are holomorphic lifts of a holomorphic map $f : U \to Z$, then there is a $g \in G$ such that $F_2 = g \circ F_1$. The same is true for local lifts of germs of holomorphic maps, and for lifts of formal morphisms.*

PROOF. To (1): Consider a local lift of f which is a germ of a holomorphic map $F : U' \to V$, where $U' \subseteq U$ is an open subset. By assumption, $F(U') \subseteq \pi^{-1}(\overline{S}) = \cup_{g \in G} gV^K$, and there is a point $x \in U'$ such that the stabilizer $G_{F(x)} = gKg^{-1}$. Then $F(U') \subseteq gV^K$ and $F' = g^{-1} \circ F$ is a local lift of f such that $F'(U') \subseteq V^K$. Then $\pi_{V^K} \circ F'$ is a germ of a holomorphic map which by construction coincides with the germ of f'. By definition, the germ f' is quasi-regular and F' is its local lift.

To (2): Let F be a formal lift of f to V. Let $I(\overline{S})$ be the prime ideal of $\mathbb{C}[Z]$ defining $\overline{S}$. Consider the pullback $\pi^*(I(\overline{S}))$ of $I(\overline{S})$. By the definition of V^K we have $\pi^{-1}(\overline{S}) = \cup_{g \in G} gV^K$. By definition of a formal lift F vanishes on $\pi^*(I(\overline{S}))$ and then, by Hilbert's Nullstellensatz, on the ideal $I(\cup_{g \in G} gV^K) = \sqrt{I(\cup_{g \in G} gV^K)}$ of $\mathbb{C}[V]$ defining the Zariski-closed subset $\cup_{g \in G} gV^K$ of V. Evidently the ideal $I(\cup_{g \in G} gV^K)$ equals the intersection of prime ideals $I(gV^K)$. Since $\mathbb{C}[[X_1, \dots, X_p]]$ is an integral domain, there is a $g \in G$ such that the formal morphism F vanishes on $I(gV^K)$ and then $F \circ g^{-1}$ is a formal lift of f which factors through the formal morphism $F' : \mathbb{C}[V^K] \to \mathbb{C}[[X_1, \dots, X_p]]$. Thus the formal morphism f factors through the formal morphism $f' = J_0^\infty(\mathbb{C}^p, \pi_{V^K})(F')$ from $\mathbb{C}^p$ to V^K/W, F' is a formal lift of f', and, by assumption, the formal morphism f' is quasi-regular.

To (3): First assume that the germ of f is quasi-regular at $x \in U$.

Let F_1 and F_2 be holomorphic lifts of f. By assumption, there is a point y in a neighborhood of x such that $F_1(y)$ and $F_2(y)$ are regular points of V. Since $(\pi \circ F_1)(y) = (\pi \circ F_2)(y)$, there is a unique $g \in G$ such that $F_2(y) = (g \circ F_1)(y)$. As the projection π is étale at $F_2(y)$, the lift F_2 coincides with $g \circ F_1$ in a neighborhood of y and then on the whole of U.

Let K be the maximal subgroup of G such that $f(U) \subseteq \pi(V^K)$, let $N_G(K)$ be the normalizer of K in G, and $W = N_G(K)/K$. By the proof of (1) the germ

$f : \mathbb{C}^p, x \to Z$ for $x \in U$ can be considered as a quasi-regular germ of a holomorphic map $f' : \mathbb{C}^p, x \to \overline{S} = \pi(V^K)$ and there are $g_1, g_2 \in G$ such that the germs $g_1 \circ F : \mathbb{C}^p, x \to V$ and $g_2 \circ F_2 : \mathbb{C}^p, x \to V$ are local lifts of the above germ f' to V^K. Then, for some $g \in H$ we have $g_2 \circ F_2 = (gg_1) \circ F_1$ in a neighborhood of x and then in the whole of U. Thus we have $F_2 = (g_2^{-1} gg_1) F_1$.

For local lifts the proof is the same. For lifts of formal morphisms the proof follows from 16.2.22 below. $\square$

The above theorem shows that the problem of lifting for local and formal lifts is reduced in some sense to one for the quasi-regular case.

Namely, let the conditions of (1) in the above theorem be satisfied. Since the morphism κ is birational, for each basic invariant τ of $\mathbb{C}[V^K]$, the composition $\kappa^{-1} \circ \tau$ is a rational function on V^K/W and, in general, the function $\kappa^{-1} \circ \tau \circ f$ is a meromorphic function on U. First we have to check that this function is analytic near x. If $f(x) \in \overline{S}_{\text{nor}}$ this is always true, because κ^{-1} is an isomorphism near $f(x)$. Then f has a local lift at x if and only if the germ $f' : \mathbb{C}^p, x \to V^K/W$ has a local lift to V^K.

The analogous statement for formal lifts is true whenever the conditions of (2) in the above theorem are satisfied.

16.2.7. Algebraic reformulation. The above geometric problem of lifting has the following algebraic interpretation. For instance, suppose that $f : \mathbb{C}^p \to Z$ is a regular morphism and F is its regular lift. Consider the morphism $f^* : \mathbb{C}[Z] = \mathbb{C}[V]^G \to \mathbb{C}[\mathbb{C}^p]$ induced by f and the morphism $F^* : \mathbb{C}[V] \to \mathbb{C}[\mathbb{C}^p]$ induced by F. Since, by definition, $\mathbb{C}[Z] \subseteq \mathbb{C}[V]$, the morphism F^* is an extension of the morphism f^* to $\mathbb{C}[V]$.

$$\begin{tikzcd} & V \arrow[d] \\ \mathbb{C}^p \arrow[ur, dashed, "F"] \arrow[r, "f"'] & Z \end{tikzcd} \qquad \begin{tikzcd} & \mathbb{C}[V] \arrow[dl, dashed, "F^*"] \\ \mathbb{C}[\mathbb{C}^p] & \mathbb{C}[Z] \arrow[l, "f^*"] \arrow[u, hook] \end{tikzcd}$$

Similarly, consider a germ of a holomorphic morphism $f : \mathbb{C}^p, 0 \to Z, O$, where $O = \pi(0)$ and its local lift $F : \mathbb{C}^p, 0 \to V, 0$. We have the morphisms $f^* : \mathfrak{F}_{Z,O} \to \mathfrak{F}_{\mathbb{C}^p,0}$ and $F^* : \mathfrak{F}_{V,0} \to \mathfrak{F}_{\mathbb{C}^p,0}$ induced by f and F respectively. Since the projection π induces the inclusion $\mathfrak{F}_{Z,O} \subseteq \mathfrak{F}_{V,0}$, the morphism F^* is an extension of the morphism f^* to $\mathfrak{F}_{V,0}$.

Finally, let $f : \mathbb{C}[Z] \to \mathbb{C}[[X_1, \ldots, X_p]]$ be a formal morphism from $\mathbb{C}^p$ to Z and let $F : \mathbb{C}[V] \to \mathbb{C}[[X_1, \ldots, X_p]]$ be its formal lift. Since the projection π induces the inclusion $\mathbb{C}[Z] \subseteq \mathbb{C}[V]$ the lift F is an extension of the morphism f to $\mathbb{C}[V]$.

16.2.8. Let τ be a homogeneous G-invariant polynomial of degree d on V and let τ^s be the corresponding symmetric d-linear form on V. For each germ $F : \mathbb{C}^p, 0 \to V$ of a holomorphic map and each system of multi-indices $(A_1, \ldots, A_d)$ we put

$$T(A_1, \ldots, A_d)(j_0^q F) := \tau^s \Big(\partial_{A_1} F(0), \ldots, \partial_{A_d} F(0) \Big).$$

By 16.2.3, $T(A_1, \ldots, A_d)$ is a function on $J_0^q(\mathbb{C}^p, V)$ for $q \geq |A_1|, \ldots, |A_d|$. From proposition 16.2.5 follows that the function

$$T(A_1, \ldots, A_d) : J_0^q(\mathbb{C}^p, V) = V^{\mathfrak{A}_{p,q}} \to \mathbb{C}$$

is regular, G-invariant, and equal to a polarization of τ up to a non-zero factor. It is also symmetric in $A_1, \ldots, A_d$.

Recall what *polarizations* are: For $v_1, \dots, v_k \in V$ and scalars $t_1, \dots, t_k$ we have

$$\begin{aligned}\tau(t_1 v_1 + \cdots + t_k v_k) =& \tau^s(t_1 v_1 + \cdots + t_k v_k, \cdots, t_1 v_1 + \cdots + t_k v_k)\\ =& \sum_{i_1,\dots,i_d} t_{i_1} \cdots t_{i_d}\, \tau^s(v_{i_1}, \dots, v_{i_d})\\ =& \sum_{r_1+\cdots+r_k=d} t_1^{r_1} \cdots t_k^{r_k}\, \tau_{r_1,\dots,r_k}(v_1, \dots, v_k),\end{aligned}$$

where

$$\tau_{r_1,\dots,r_k}(v_1,\dots,v_k) := \frac{d!}{r_1!\dots r_k!}\, \tau^s(\underbrace{v_1,\dots,v_1}_{r_1 \text{ times}}, \dots, \underbrace{v_k,\dots,v_k}_{r_k \text{ times}}).$$

The polynomials $\tau_{r_1,\dots,r_k}$ are the usual polarizations of τ on V^k.

By proposition 16.2.4, we may define a rational function $\tilde{T}(A_1, \dots, A_d)$ on $\bar{J}_0^q(\mathbb{C}^p, Z)$ by the condition $T(A_1, \dots, A_d) = \tilde{T}(A_1, \dots, A_d) \circ J_0^q(\mathbb{C}^p, \pi)$. By definition, we have

$$\tilde{T}(\emptyset, \dots, \emptyset) \circ \pi = T(\emptyset, \dots, \emptyset) = \tau. \tag{16.1}$$

Now extend the d-linear form τ^s on V to a d-linear form $\mathfrak{T}$ on $J_0^\infty(\mathbb{C}^p, V) = V \otimes \mathbb{C}[[X_1, \dots, X_p]]$ with values in $\mathbb{C}[[X_1, \dots, X_p]]$ which is defined by the following condition. For $1 \le i \le d$, $v_i \in V$, and $F_i \in \mathbb{C}[[X_1, \dots, X_p]]$

$$\mathfrak{T}(v_1 \otimes F_1, \dots, v_d \otimes F_d) := \tau^s(v_1, \dots, v_d)\, F_1 \cdots F_d.$$

For $h \in J_0^\infty(\mathbb{C}^p, V) = V \otimes \mathbb{C}[[X_1, \dots, X_p]]$ and a system of multi-indices $A_1, \dots, A_d$ put

$$\mathfrak{T}(A_1, \dots, A_d)(h) := \mathfrak{T}(\partial_{A_1} h, \dots, \partial_{A_d} h).$$

By definition, the function $\mathfrak{T}(A_1, \dots, A_d) : J_0^\infty(\mathbb{C}^p, V) \to \mathbb{C}[[X_1, \dots, X_p]]$ is G-invariant and symmetric in $A_1, \dots, A_d$.

$$\begin{tikzcd} J_0^\infty(\mathbb{C}^p, V) \arrow[r, equal] & V \otimes \mathbb{C}[[X_1, \dots, X_p]] \arrow[r, "\mathfrak{T}(A_1,\dots,A_d)"] & \mathbb{C}[[X_1, \dots, X_p]] \end{tikzcd}$$

$$\begin{tikzcd} J_0^q(\mathbb{C}^p, V) \arrow[r, "T(A_1,\dots,A_d)"] \arrow[d, "J_0^q(\mathbb{C}^p,\pi)"'] & \mathbb{C} \\ \bar{J}_0^q(\mathbb{C}^p, Z) \arrow[ur, "\tilde{T}(A_1,\dots,A_d)"'] & \end{tikzcd}$$

16.2.9. The q-jet of the identity map on V in invariant coordinates. Let $v : V \to V$ be the identity map. Let $v_0 \in V_0$ be a regular point of V and let y_i be invariant coordinates in a neighborhood U of v_0 in V introduced in 16.2.2. Then in U the map v is defined by a holomorphic function $v(y_i)$ with values in V.

Let $I = (i_1, \dots, i_s)$ be a (unordered) multi-index with $i_1, \dots, i_s \in \{1, \dots, n\}$. In particular, for $s = 0$ we put $I := \emptyset$. Then the q-jet of the identity map v at each point $x \in U$ is defined by the set of partial derivatives $\partial_I v = \frac{\partial^s v}{\partial y_{i_1} \dots \partial y_{i_s}}$ for $|I| \le q$ at x.

Let e_a be a basis of V, let u_a be the corresponding coordinates, and let $J = \det(\frac{\partial y_i}{\partial u_j})$ be the Jacobian.

LEMMA. *Let $I = (i_1, \dots, i_s)$, where $s > 0$, be a multi-index. Then $\widetilde{\partial_I v} := J^{2s-1} \partial_I v$ is a regular map from U to V.*

PROOF. We prove this lemma by induction with respect to s. We use that

$$\frac{\partial u_a}{\partial y_i} = \frac{1}{J} J_i^a,$$

where $J_i^a \in \mathbb{C}[V]$ is the cofactor of the entry $\frac{\partial y_i}{\partial u_a}$ in the Jacobi matrix $(\frac{\partial y_i}{\partial u_j})$.

For $s = 1$, $J\partial_i v = \sum_{a=1}^n J_i^a e_a$ is a regular map from U to V.

Let $I = (i_1, \ldots, i_s)$ with $\widetilde{\partial_I v}$ being regular. Then for $I' = (i_1, \ldots, i_{s+1})$ we have

$$\partial_{I'} v = \frac{\partial}{\partial y_{i_{s+1}}} \Big(\frac{\widetilde{\partial_I v}}{J^{2s-1}} \Big) = \frac{1}{J^{2s+1}} \sum_{a=1}^n J_{i_{s+1}}^a \Big(J \frac{\partial(\widetilde{\partial_I v})}{\partial u_a} - (2s-1) \frac{\partial J}{\partial u_a} \widetilde{\partial_I v} \Big),$$

where $\sum_{a=1}^n J_{i_{s+1}}^a \Big(J \frac{\partial(\widetilde{\partial_I v})}{\partial u_a} - (2s-1) \frac{\partial J}{\partial u_a} \widetilde{\partial_I v} \Big)$ is a regular map from U to V. $\square$

16.2.10. Let $h = j_0^q F \in J_0^q(\mathbb{C}^p, V)$, where $F : \mathbb{C}^p, 0 \to V$ is a germ of a holomorphic map such that $F(0) \in V_0$. Put $F_i := y_i \circ F$, where y_i are the invariant coordinates on V. We need to express the q-jet $j_0^q F$ in terms of the q-jet of the identity map v, i.e., we have to find the explicit formula for each $h_A = \partial_A F(0)$ with $A \in \mathfrak{A}_{p,q}$ in terms of $\partial_B F_i$ and $\partial_I v$ with $|B|, |I| \leq |A|$. We can extract this formula from the following expression (see the classical Faa di Bruno formula for $p = 1$ and [**Ver83**] for arbitrary p).

$$d^q(v \circ F) = q! \sum_{k=1}^q \frac{1}{k!} \sum_{i_1,\ldots,i_k=1}^n \Big(\frac{\partial^k v}{\partial y_{i_1} \ldots \partial y_{i_k}} \circ F \Big) \sum_{q_1+\cdots+q_k=q} \frac{d^{q_1} F_{i_1}}{q_1!} \cdots \frac{d^{q_k} F_{i_k}}{q_k!}. \tag{16.2}$$

where y_i are arbitrary local coordinates in V. Note that the formula (16.2) is true whenever $F : \mathbb{C}[V] \to \mathbb{C}[[X_1, \ldots, X_p]]$ is a formal morphism from $\mathbb{C}^p$ to V and $F_i = F(y_i)$. It implies the following

LEMMA. *For each multi-index $A = (a_1, \ldots, a_s) \neq \emptyset$, where $a_1, \ldots, a_s \in \{1, \ldots, p\}$, there is a well defined function*

$$\Psi_A : (X, Y) \mapsto \sum_{1 \leq |I| \leq |A|} a_{A,I}(Y)\, X_I,$$

where $X = (X_I)_{1 \leq |I| \leq |A|}$ and where the coefficients $a_{A,I}$ are polynomials in $Y = (y_{i,B})_{1 \leq i \leq n, 1 \leq |B| \leq |A|}$, such that for each germ of a holomorphic map $F : \mathbb{C}^p, 0 \to V, v$ with v regular and for the local coordinates y_i from above we have

$$\partial_A F = \Psi_A \Big((\partial_I v \circ F)_I, (\partial_B F_i)_{i,B} \Big).$$

16.2.11. The function $\tilde{T}(A_1, \ldots, A_d)$. We consider $T(A_1, \ldots, A_d)$ which is a regular function on $J_0^q(\mathbb{C}^p, V)$, and $\tilde{T}(A_1, \ldots, A_d)$ which is a rational function on $\bar{J}_0^q(\mathbb{C}^p, Z)$, both defined in 16.2.8.

Let z_i be the regular function on Z for $1 \leq i \leq n$, used in 16.2.2 for the construction of the invariant coordinates on V. Let $A_1, \ldots, A_{d'} \neq \emptyset$ and $A_{d'+1} = \cdots = A_d = \emptyset$. Put $M := 2(|A_1| + \cdots + |A_d|) - d'$. By lemma 16.2.9, for any system of multi-indices $I_1, \ldots, I_d$ such that $1 \leq |I_1| \leq |A_1|, \ldots, 1 \leq |I_{d'}| \leq |A_{d'}|$ and $I_{d'+1} = \cdots = I_d = \emptyset$, the expression $\Delta_M \cdot T \circ (\partial_{I_1} v, \ldots, \partial_{I_d} v)$ is a G-invariant and regular function on V. Thus there is a unique rational function $\tilde{T}(I_1, \ldots, I_d)$ on Z such that $\tilde{T}(I_1, \ldots, I_d) \circ \pi = T(\partial_{I_1} v, \ldots, \partial_{I_d} v)$ and $\tilde{\Delta}_M \cdot \tilde{T}(I_1, \ldots, I_d)$ is a regular function on Z.

Let q be the maximal order of the multi-indices $A_1, \ldots, A_d$. For $k = 1, \ldots, d'$ we may consider the a_{A_k, I_k} of 16.2.10 as polynomials in $Y = (y_{i,B})_{1 \leq i \leq n, 1 \leq |B| \leq q}$. Put

$$a_{A_1, \ldots, A_{d'}, I_1, \ldots, I_{d'}}(Y) := a_{A_1, I_1}(Y) \cdots a_{A_{d'}, I_{d'}}(Y).$$

THEOREM. *Let $A_1, \dots, A_{d'} \neq \emptyset$, $A_{d'+1}, \dots, A_d = \emptyset$ and $A_1, \dots, A_d \in \mathfrak{A}_{p,q}$. Then:*

(1) *The following is a rational function $\tilde{T}(A_1, \dots, A_d)$ on $\bar{J}^q_0(\mathbb{C}^p, Z)$:*

$$\tilde{T}(A_1, \dots, A_d) := \sum_{\substack{1 \le |I_1| \le |A_1|, \\ \vdots \\ 1 \le |I_{d'}| \le |A_{d'}|}} a_{A_1,\dots,A_{d'},I_1,\dots,I_{d'}} \Big((\partial_B z_i)_{i,B} \Big) \, \tilde{T}(I_1, \dots, I_{d'}, \emptyset, \dots, \emptyset).$$

(2) $\tilde{T}(A_1, \dots, A_d) \circ J^q_0(\mathbb{C}^p, \pi) = T(A_1, \dots, A_d)$.

(3) $\tilde{\Delta}_M \cdot \tilde{T}(A_1, \dots, A_d)$ *is a regular function on* $J^q_0(\mathbb{C}^p, Z)$.

$$\begin{tikzcd} J^q_0(\mathbb{C}^p, V) \arrow[r, "T(A_1,\dots,A_d)"] \arrow[d, "J^q_0(\mathbb{C}^p,\pi)"'] & \mathbb{C} \\ \bar{J}^q_0(\mathbb{C}^p, Z) \arrow[ur, dashed, "\tilde{T}(A_1,\dots,A_d)"'] & \end{tikzcd}$$

PROOF. To (1): By proposition 16.2.4, it suffices to check that the condition $\tilde{T}(A_1, \dots, A_d) \circ \pi^q = T(A_1, \dots, A_d)$ is satisfied for the above expression of $\tilde{T}(A_1, \dots, A_d)$. By lemma 16.2.10, this follows from

$$T(A_1, \dots, A_d)(h) = \sum_{\substack{1 \le |I_1| \le |A_1|, \\ \vdots \\ 1 \le |I_{d'}| \le |A_{d'}|}} \Big(a_{A_1,\dots,A_{d'},I_1,\dots,I_{d'}} \left((\partial_B F_i)_{i,B} \right) T(\partial_{I_1} v, \dots, \partial_{I_{d'}} v, v, \dots, v) \circ F \Big)(0), \tag{16.3}$$

where $h = j^q_0 F \in J^q_0(\mathbb{C}^p, V)$.

To (2): This statement follows from proposition 16.2.4.

To (3): This statement follows from (1) and lemma 16.2.9. □

16.2.12. Let $F \in V \otimes \mathbb{C}[[X_1, \dots, X_p]]$ be a formal morphism from $\mathbb{C}^p$ to V and $F_i = y_i \circ F$. Lemma 16.2.10 implies the following

LEMMA. *For each multi-index A such that $|A| \geq 1$ and the invariant coordinates y_i on V we have*

$$J^{2|A|-1}(F) \, \partial_A F = \Psi_A \bigg(\Big((J^{2(|A|-|I|)} \widetilde{\partial_I v})(F) \Big)_I, \Big(\partial_B F_i \Big)_{i,B} \bigg).$$

□

16.2.13. The function $\tilde{\mathfrak{T}}(A_1, \dots, A_d)$. Consider the presentation of Z as an irreducible Zariski-closed subset of $\mathbb{C}^m$ defined in 16.2.1. Denote by $I(Z)$ the prime ideal of the ring of polynomials $\mathbb{C}[W_1, \dots, W_m]$ defining $Z \subseteq \mathbb{C}^m$. Each formal morphism $f \in J^\infty_0(\mathbb{C}^p, Z)$ is defined by the equations $f(W_j) = f_j$ for $1 \le j \le m$, where $f_j \in \mathbb{C}[[X_1, \dots, X_p]]$ and $\Phi(f_j) = 0$ for each $\Phi \in I(Z)$.

Let ψ be a regular function on Z which is the restriction to Z of a polynomial $\Psi \in \mathbb{C}[W_1, \dots, W_m]$. For each $f \in J^\infty_0(\mathbb{C}^p, Z)$ put $\psi(f) := \Psi(f_j)$. By definition we have $\psi(f) = f(\psi)$, where f is considered as a morphism $\mathbb{C}[Z] \to \mathbb{C}[[X_1, \dots, X_p]]$. Then $\psi(f)$ defines a unique function $J^\infty_0(\mathbb{C}^p, Z) \to \mathbb{C}[[X_1, \dots, X_p]]$ which is independent of the choice of the polynomial Ψ.

Similarly, consider a rational function ψ on Z such that $\psi = \frac{\psi_1}{\psi_2}$, where ψ_1 and ψ_2 are regular functions on Z and put $\psi(f) := \frac{\psi_1(f)}{\psi_2(f)}$ whenever $\psi_2(f) \neq 0$. It is clear that $\psi(f)$ is a function on $J^\infty_0(\mathbb{C}^p, Z)$ with values in the field $\mathbb{C}((X_1, \dots, X_p))$

of fractions of the ring $\mathbb{C}[[X_1,\dots,X_p]]$ which is independent of the choice of the presentation $\psi = \frac{\psi_1}{\psi_2}$.

Let z_i be the regular functions on Z used in 16.2.2 for the construction of the invariant coordinates y_i on V. Let $f \in J_0^\infty(\mathbb{C}^p, Z)$ be a quasi-regular formal morphism from $\mathbb{C}^p$ to Z such that $\tilde{\Delta}(f) \neq 0$. For $A_1,\dots,A_{d'} \neq \emptyset$, $A_{d'+1},\dots,A_d = \emptyset$ and $M = 2(|A_1| + \dots + |A_d|) - d'$ put

$$\tilde{\mathfrak{T}}(A_1,\dots,A_d)(f) :=$$
$$= \sum_{\substack{1\le|I_1|\le|A_1|,\\ \dots \\ 1\le|I_{d'}|\le|A_{d'}|}} \Big(a_{A_1,\dots,A_{d'},I_1,\dots,I_{d'}}\big((\partial_B z_i)_{i,B}\big)\cdot \tilde{T}(I_1,\dots,I_{d'},\emptyset,\dots,\emptyset)\Big)(f). \quad (16.4)$$

By definition, $\tilde{\mathfrak{T}}(A_1,\dots,A_d)$ is a function with values in the field $\mathbb{C}((X_1,\dots,X_p))$ on the set $\bar{J}_0^\infty(\mathbb{C}^p, Z)$ consisting of all $f \in J_0^\infty(\mathbb{C}^p, Z)$ such that $\tilde{\Delta}(f) \neq 0$.

THEOREM. *Let $A_1,\dots,A_{d'} \neq \emptyset$, $A_{d'+1},\dots,A_d = \emptyset$, and $M = 2(|A_1| + \dots + |A_d|) - d'$. Then the function $\tilde{\mathfrak{T}}(A_1,\dots,A_d)$ satisfies the following conditions:*

(1) *$\tilde{\mathfrak{T}}(A_1,\dots,A_d) \circ J_0^\infty(\mathbb{C}^p,\pi) = \mathfrak{T}(A_1,\dots,A_d)$, where $\mathfrak{T}$ is from 16.2.8.*
(2) *$\tilde{\Delta}_M \cdot \tilde{\mathfrak{T}}(A_1,\dots,A_d)$, where $\tilde{\Delta}_M$ is regarded as a function on $J_0^\infty(\mathbb{C}^p, Z)$, is a function on $J_0^\infty(\mathbb{C}^p, Z)$ with values in $\mathbb{C}[[X_1,\dots,X_p]]$.*

$$\begin{tikzcd}
J_0^\infty(\mathbb{C}^p, V) \arrow[r, "\mathfrak{T}(A_1,\dots,A_d)"] \arrow[d, "J_0^\infty(\mathbb{C}^p,\pi)"'] & \mathbb{C}[[X_1,\dots,X_p]] \\
\bar{J}_0^\infty(\mathbb{C}^p, Z) \arrow[ur, dashed, "\tilde{\mathfrak{T}}(A_1,\dots,A_d)"'] &
\end{tikzcd}$$

PROOF. The proof follows from the definition of the function $\tilde{\mathfrak{T}}(A_1,\dots,A_d)$ and lemma 16.2.12. ☐

16.2.14. Local lifts at regular points.

PROPOSITION. *Let $f : \mathbb{C}^p, x \to Z, z$ be a germ at $x \in \mathbb{C}^p$ of a holomorphic map with z regular. Then for each $v \in \pi^{-1}(z)$ there is a unique local holomorphic lift $F : \mathbb{C}^p, x \to V, v$ of f.*

PROOF. This follows from proposition 16.2.1. ☐

16.2.15. Lifts of quasi-regular holomorphic germs. Let X be an affine variety and let f be either a rational morphism from $\mathbb{C}^p$ to X or a holomorphic map defined on a classically open connected subset $U \subseteq \mathbb{C}^p$ to X. Consider the morphism $j^q f$ from $\mathbb{C}^p$ or from U to $J_0^q(\mathbb{C}^p, X)$, which for $x \in U$ is given by $j^q f(x) = j_0^q f(\quad + x)$. The morphism $j^q f$ is rational and is regular wherever f is regular; or holomorphic if f is holomorphic.

Let $\sigma : V \to \sigma(V) \subseteq \mathbb{C}^m$ be the morphism defined by the system of basic generators $\sigma_1,\dots,\sigma_m$. Recall that $\sigma(V)$ and $Z = V/G$ are isomorphic as affine varieties and, for this presentation of Z, the maps σ and $\pi : V \to V/G$ coincide.

Denote by $w_1,\dots,w_m$ the standard coordinates in $\mathbb{C}^m$ and let $I(Z)$ be the prime ideal of the ring $\mathbb{C}[W_1,\dots,W_m]$ defining Z. Consider $\mathbb{C}[W_1,\dots,W_m]$ as a graded ring with a grading defined by $\deg W_j = \deg \sigma_j$ for $1 \le j \le m$. Then $I(Z)$ is a homogeneous ideal.

For $\tau = \sigma_j$ and a system $A_1,\dots,A_{d_j}$ of multi-indices denote by $\tilde{S}_j(A_1,\dots,A_{d_j})$ the rational function $\tilde{T}(A_1,\dots,A_{d_j})$ from 16.2.8 on $\bar{J}_0^q(\mathbb{C}^p, Z)$.

Recall that, by 16.2.6, for a quasi-regular germ $f : \mathbb{C}^p, 0 \to Z$ of a holomorphic map there is a choice of invariant coordinates such that for the corresponding function $\tilde{\Delta}$ we have $\tilde{\Delta} \circ f \neq 0$.

THEOREM. *Consider a quasi-regular germ $f : \mathbb{C}^p, 0 \to Z = V/G$ of a holomorphic map described by $w_j \circ f = f_j$ for $1 \leq j \leq m$. Assume that, for some choice of invariant coordinates such that $\tilde{\Delta} \circ f \neq 0$, q is the minimal order of non-zero terms of the Taylor expansion of $\tilde{\Delta} \circ f$ at 0.*

Then the lift F of f at 0 exists if and only if for $1 \leq j \leq m$ and for each system of multi-indices $A_1, \ldots, A_{d_j} \in \mathfrak{A}_{p,q}$ the functions $\tilde{S}_j(A_1, \ldots, A_{d_j}) \circ j^q f$ have holomorphic extensions to a neighborhood of 0.

$$\begin{tikzcd}
 & J_0^q(\mathbb{C}^p, V) \arrow[r, "S(A_1,\ldots,A_d)"] \arrow[d, "J_0^q(\mathbb{C}^p,\pi)"'] & \mathbb{C} \\
\mathbb{C}^p, 0 \arrow[ur, "j^q F"] \arrow[r, "j^q f"] & \bar{J}_0^q(\mathbb{C}^p, Z) \arrow[ur, "\tilde{S}(A_1,\ldots,A_d)"'] &
\end{tikzcd}$$

PROOF. Let F be a lift of f. Then by the definition of the function $\tilde{S}_j(A_1, \ldots, A_{d_j})$ for each $q \geq 0$ we have

$$\tilde{S}_j(A_1, \ldots, A_{d_j}) \circ j^q f = S_j(A_1, \ldots, A_{d_j}) \circ j^q F : \mathbb{C}^p, 0 \to \mathbb{C},$$

where the right hand side defines a holomorphic germ.

Conversely, let the assumptions of the theorem be satisfied. Let us now use a representative $f : U \to Z$ of the germ, where U is a connected open neighborhood of 0. Let q be the minimal order of non-zero terms of the Taylor expansion of $\tilde{\Delta} \circ f$ at 0. For each $1 \leq j \leq m$ consider the function

$$f_j^q(x, t) := \sum_{A_1, \ldots, A_{d_j} \in \mathfrak{A}_{p,q}} \Big(\tilde{S}_j(A_1, \ldots, A_{d_j}) \circ j^q f \Big)(x)\, t_{A_1} \cdots t_{A_{d_j}}, \tag{16.5}$$

where $x \in U$ and $t = (t_A)_{A \in \mathfrak{A}_{p,q}} \in \mathbb{C}^{\mathfrak{A}_{p,q}}$. By assumption, the function f_j^q is a polynomial in t whose coefficients are holomorphic near 0. By definition, the map $f^q = (f_1^q, \ldots, f_m^q) : \mathbb{C}^p \times \mathbb{C}^{\mathfrak{A}_{p,q}} \to \mathbb{C}^m$ is holomorphic near 0.

Since $Z_{>0}$ is a Zariski-closed subset of Z of codimension ≥ 1, the inverse image $f^{-1}(Z_{>0})$ is a complex analytic subset of U of codimension ≥ 1 and $f^{-1}(Z_0)$ is a dense open subset of U, since f is quasi-regular.

Let, for $y \in U$, $f(y)$ be a regular point and let F_y be a local lift of f defined in a neighborhood U_y of y, which exists by proposition 16.2.14. For each q consider the holomorphic map $F_y^q : U_y \times \mathbb{C}^{\mathfrak{A}_{p,q}} \to V$ given by:

$$F_y^q(x, t) := \sum_{A \in \mathfrak{A}_{p,q}} \partial_A F_y(x)\, t_A. \tag{16.6}$$

By theorem 16.2.11, we have

$$\begin{aligned}
(\sigma_j \circ F_y^q)(x, t) &= \sum_{A_1, \ldots, A_{d_j} \in \mathfrak{A}_{p,q}} S_j\Big(\partial_{A_1} F_y(x), \ldots, \partial_{A_{d_j}} F_y(x)\Big)\, t_{A_1} \cdots t_{A_{d_j}} \\
&= \sum_{A_1, \ldots, A_{d_j} \in \mathfrak{A}_{p,q}} \Big(S_j(A_1, \ldots, A_{d_j}) \circ j^q F_y \Big)(x)\, t_{A_1} \cdots t_{A_{d_j}} = f_j^q(x, t).
\end{aligned} \tag{16.7}$$

Therefore for each polynomial $\Phi \in I(Z)$ we have $\Phi \circ f^q = 0$ on $U_y \times \mathbb{C}^{\mathfrak{A}_{p,q}}$ and thus also on $U \times \mathbb{C}^{\mathfrak{A}_{p,q}}$. So f^q is a holomorphic map from $U \times \mathbb{C}^{\mathfrak{A}_{p,q}}$ to Z and F_y^q is a lift of f^q.

For each germ of a holomorphic function $\varphi \in \mathfrak{F}_{\mathbb{C}^p,x}$, denote by $\operatorname{Tay}_x^q \varphi$ the sum of terms of the Taylor expansion at x of φ of orders $\leq q$. For each germ $\varphi = (\varphi_j)_j$ of a holomorphic map $\mathbb{C}^p, x \to \mathbb{C}^m$, put $\operatorname{Tay}_x^q \varphi := (\operatorname{Tay}_x^q \varphi_j)_j$.

By assumption, there is a multi-index $A \in \mathfrak{A}_{p,q}$ such that $\partial_A(\tilde{\Delta} \circ f)(0) = \partial_A(\tilde{\Delta} \circ \operatorname{Tay}_0^q f)(0) \neq 0$. This implies that there is a point $x_0 = (x_{0,1}, \dots, x_{0,p}) \in \mathbb{C}^p$ such that $(\tilde{\Delta} \circ \operatorname{Tay}_0^q f)(x_0) \neq 0$.

For

$$A = (\underbrace{1, \dots, 1}_{r_1 \text{ times}}, \dots, \underbrace{p, \dots, p}_{r_p \text{ times}}),$$

put

$$t_A(x) := \frac{1}{r_1! \dots r_p!} (x_1)^{r_1} \dots (x_p)^{r_p}, \quad t(x) := (t_A(x))_{A \in \mathfrak{A}_{p,q}}$$

where $x = (x_1, \dots, x_p) \in \mathbb{C}^p$.

By definition, we have $F_y^q(y, t(x-y)) = \operatorname{Tay}_y^q F_y(x)$ and then $f_j^q(y, t(x-y)) = (\sigma_j \circ \operatorname{Tay}_y^q F_y)(x)$. On the other hand, since σ_j is homogeneous, for a fixed y we have $\operatorname{Tay}_y^q f_j = \operatorname{Tay}_y^q(\sigma_j \circ F_y) = \operatorname{Tay}_y^q(\sigma_j \circ \operatorname{Tay}_y^q F_y)$. Thus, we have

$$\operatorname{Tay}_y^q f_j^q(y, t(x-y)) = \operatorname{Tay}_y^q f_j(x). \tag{16.8}$$

By assumption, the function $\tilde{S}_j(A_1, \dots, A_{d_j}) \circ j^q f$ has a holomorphic extension to a neighborhood of 0 and we may suppose that the point y belongs to this neighborhood. Letting $y \to 0$ in (16.8), we get $\operatorname{Tay}_0^q f_j^q(0, t(x)) = \operatorname{Tay}_0^q f_j(x)$. Then we have $(\tilde{\Delta} \circ f^q)(0, t(x)) = \tilde{\Delta}(\operatorname{Tay}_0^q f)(x)$, and, for the point $x_0 \in \mathbb{C}^p$ chosen above, we have $(\tilde{\Delta} \circ f^q)(0, t(x_0)) \neq 0$, i.e., $f^q(0, t(x_0))$ is a regular point of Z.

Now we will construct a local lift of f. Consider a local holomorphic lift F^q of f^q near $(0, t(x_0))$ in $U \times \mathbb{C}^{\mathfrak{A}_{p,q}}$ which exists by proposition 16.2.14. We can choose y near 0 so that $f(y) \in Z_0$, so there exists a local holomorphic lift F_y of f near y, and still $(\tilde{\Delta} \circ f^q)(y, t(x_0)) \neq 0$. Consider the map F_y^q defined by formula (16.6). Both F_y^q and F^q are local lifts of f^q at $(y, t(x_0))$. By theorem 16.2.6, there exists $g \in G$ such that $F_y^q = gF^q$ near $(y, t(x_0))$.

Since $F_y^q(x,t)$ is linear in $t \in \mathbb{C}^{\mathfrak{A}_{p,q}}$, also $F^q(x,t)$ is linear in t and thus is defined for all t. Put $t_1 := (t_{1,A})_A$, where $t_{1,\emptyset} = 1$ and $t_{1,A} = 0$ for $A \neq \emptyset$. Then near $0 \in \mathbb{C}^p$ we have by (16.1),

$$\sigma_j(F^q(x, t_1)) = f_j^q(x, t_1) = \left(\tilde{S}_j(\emptyset, \dots, \emptyset) \circ j^q f\right)(x) = f_j(x),$$

i.e., $F^q(\quad, t_1)$ is a local lift of f at 0. $\square$

Remark. Consider the grading of the ring $\mathbb{C}[Z] = \mathbb{C}[V]^G$ induced by the natural grading of the polynomial ring $\mathbb{C}[V]$ and denote by r the order of the homogeneous function $\tilde{\Delta}$. Let $f : \mathbb{C}^p, 0 \to Z$ be a germ of a holomorphic map satisfying for some positive integer q and for each $1 \leq j \leq m$ the following conditions:

(1) The function $\tilde{S}_j(A_1, \dots, A_{d_j}) \circ j^q f$ has a holomorphic extension to a neighborhood of 0 for each system of multi-indices $A_1, \dots, A_{d_j} \in \mathfrak{A}_{p,q}$ such that, $\left(\tilde{S}_j(A_1, \dots, A_{d_j}) \circ j^{q-1} f\right)(0) = 0$ for all $A_1, \dots, A_{d_j} \in \mathfrak{A}_{p,q-1}$;

(2) $\operatorname{Tay}_0^{rq}(\tilde{\Delta}(f)) \neq 0$.

Then f has a local lift at 0.

Actually, since $\operatorname{Tay}_0^{rq}(\tilde{\Delta}(f)) = \operatorname{Tay}_0^{rq}(\tilde{\Delta}(\operatorname{Tay}_0^q f))$, the proof of theorem 16.2.15 is valid for this q.

16.2.16. Global holomorphic lifts. The following theorem shows that the problem of global holomorphic lifting can be described topologically.

THEOREM. *Let $U \subseteq \mathbb{C}^p$ be a classically open connected subset of $\mathbb{C}^p$ and let $f : U \to Z = V/G$ be a holomorphic map such that $f^{-1}(Z_0) \neq \emptyset$. Then a holomorphic lift $F : U \to V$ exists if and only if the image of the fundamental group $\pi_1(f^{-1}(Z_0))$ under f is contained in the image of the fundamental group $\pi_1(V_0)$ under the projection π.*

PROOF. Since by proposition 16.2.14, a local holomorphic lift of f exists for each $x \in f^{-1}(Z_0)$, the condition of the theorem is equivalent to the existence of a holomorphic lift for the restriction of f to $f^{-1}(Z_0)$. Actually, let F be such a lift. Since $f^{-1}(Z_0)$ is an open dense subset of U and π is a finite morphism, the lift F is bounded on bounded subsets of $U \cap f^{-1}(Z_0)$. Then by Riemann's extension theorem F has a holomorphic extension to U which is a holomorphic lift of f. $\square$

16.2.17. The problem of the existence of a global regular lift reduces to the one for a holomorphic lift.

THEOREM. *Let $U \subseteq \mathbb{C}^p$ be a Zariski-open subset of $\mathbb{C}^p$ and let $f : \mathbb{C}^p \to Z = V/G$ be a rational morphism which is regular in U and such that $f^{-1}(Z_0) \neq \emptyset$. If a global holomorphic lift of f on U exists then it is regular.*

PROOF. The proof follows from [**LMP03**, Lemma 5.1.1]. $\square$

16.2.18. Global regular lifts.

THEOREM. *Let $f : \mathbb{C}^p \to Z = V/G$ be a regular morphism such that $f^{-1}(Z_0) \neq \emptyset$. Then f has a regular lift if and only if there is an integer $q > 0$ such that, for each $1 \leq j \leq m$ and each multi-index $A = (a_1, \ldots, a_q)$, the mapping $\tilde{S}_j(A, \ldots, A) \circ j^q f$ is constant.*

PROOF. Let u_i be linear coordinates in V, and let $F = (F_1, \ldots, F_n)$ be the expression of a regular lift of f in these coordinates. Suppose q is the maximal degree of the polynomials F_i. Then for each $A = (a_1, \ldots, a_q)$ and each j the map $S_j(A, \ldots, A) \circ j^q F$ is constant. Theorem 16.2.11 implies that $\tilde{S}_j(A, \ldots, A) \circ j^q f$ is constant as well.

Let the condition of the theorem be satisfied. Let $x \in \mathbb{C}^p$ be a point such that $f(x) \in Z_0$. Then there is a local lift F of f at x. By theorem 16.2.15, the condition of the theorem implies that $\sigma_j(\partial_A F)$ is constant for each $A = (a_1, \ldots, a_q)$ and each j. But this means that $\partial_A F$ is constant also and, therefore, F is a polynomial map of degree $\leq q$ in a neighborhood of x. Thus F has a polynomial extension to the whole of $\mathbb{C}^p$ and this extension is a lift of f. $\square$

16.2.19. Local lifts at regular points. In the following special case the existence of local lifts implies the existence of a global lift.

THEOREM. *Let $f = (f_j) : \mathbb{C}^p \to \mathbb{C}^m$ be a regular morphism from $\mathbb{C}^p$ to $Z = V/G$ such that each function f_j is homogeneous of degree rd_j for some positive integer r and $d_j = \deg \sigma_j$. Then a global regular lift of f exists if and only if f has a local holomorphic lift at $0 \in \mathbb{C}^p$.*

PROOF. Consider the action of the group $\mathbb{C}^*$ on Z induced by the action of the homothety group on V. This action induces a homotopy equivalence between the open subset $f^{-1}(Z_0)$ and the open subset $f^{-1}(Z_0) \cap B$, where B is an open ball in $\mathbb{C}^p$ centered at 0. Then the statement of the theorem follows from theorems 16.2.16 and 16.2.17. $\square$

16.2.20. The functions $P_q(\tau)$ and $\tilde{P}_q(\tau)$. Proposition 16.2.14 about the existence of lifts at regular points also holds for formal morphisms.

For a homogeneous G-invariant polynomial τ on V, consider the function $P_q(\tau) : J_0^\infty(\mathbb{C}^p, V) \to \mathbb{C}[(t_A)_A] \otimes \mathbb{C}[[X_1, \dots, X_p]]$ and the function $\tilde{P}_q(\tau)$ on the set of quasi-regular formal morphisms $f \in J_0^\infty(\mathbb{C}^p, Z)$ such that $\tilde{\Delta}(f) \neq 0$ with values in $\mathbb{C}[(t_A)_A] \otimes \mathbb{C}((X_1, \dots, X_p))$, where $(t_A)_A = (t_A)_{A \in \mathfrak{A}_{p,q}}$ and $\mathbb{C}[(t_A)_A]$ is the ring of polynomials in $(t_A)_A$ with complex coefficients, defined as follows:

$$P_q(\tau)(F) := \sum_{A_1, \dots, A_d \in \mathfrak{A}_{p,q}} \mathfrak{T}(A_1, \dots, A_d)(F)\, t_{A_1} \dots t_{A_d}, \tag{16.9}$$

$$\tilde{P}_q(\tau)(f) := \sum_{A_1, \dots, A_d \in \mathfrak{A}_{p,q}} \tilde{\mathfrak{T}}(A_1, \dots, A_d)(f)\, t_{A_1} \dots t_{A_d}, \tag{16.10}$$

where $F \in J_0^\infty(\mathbb{C}^p, V)$, $f \in J_0^\infty(\mathbb{C}^p, Z)$, and where $\mathfrak{T}(A_1, \dots, A_d)$ and $\tilde{\mathfrak{T}}(A_1, \dots, A_d)$ are the functions defined in 16.2.8 and 16.2.13.

The following lemma follows from the definitions of $P_q(\tau)$, $\tilde{P}_q(\tau)$, $\tilde{\mathfrak{T}}(A_1, \dots, A_d)$, and theorem 16.2.13.

Lemma. (1) *We have*

$$P_q(\tau) = \tilde{P}_q(\tau) \circ J_0^\infty(\mathbb{C}^p, \pi) : J_0^\infty(\mathbb{C}^p, V) \to \mathbb{C}[(t_A)_A] \otimes \mathbb{C}[[X_1, \dots, X_p]].$$

$$\begin{tikzcd} J_0^\infty(\mathbb{C}^p, V) \arrow[r, "P_q(\tau)"] \arrow[d, "J_0^\infty(\mathbb{C}^p,\pi)"'] & \mathbb{C}[(t_A)_A] \otimes \mathbb{C}[[X_1, \dots, X_p]] \\ \bar{J}_0^\infty(\mathbb{C}^p, Z) \arrow[ur, "\tilde{P}_q(\tau)"'] & \end{tikzcd}$$

(2) *If $\tau_1, \tau_2 \in \mathbb{C}[V]^G$ are homogeneous polynomials of the same degree, then*

$$\tilde{P}_q(\tau_1 + \tau_2) = \tilde{P}_q(\tau_1) + \tilde{P}_q(\tau_2).$$

(3) *Let $\tau_1, \tau_2 \in \mathbb{C}[V]^G$ be homogeneous polynomials. Then we have*

$$\tilde{P}_q(\tau_1 \tau_2) = \tilde{P}_q(\tau_1)\tilde{P}_q(\tau_2).$$

(4) *Let f be a polynomial in the graded variables $T_1, \dots, T_r$ of degrees $d_1, \dots, d_r$ which is homogeneous with respect to this grading, and let $\tau_1, \dots, \tau_r \in \mathbb{C}[V]^G$ be homogeneous polynomials of degrees $d_1, \dots, d_r$. Then we have*

$$P_q\Big(f(\tau_1, \dots, \tau_r)\Big) = f\Big(P_q(\tau_1), \dots, P_q(\tau_r)\Big),$$
$$\tilde{P}_q\Big(f(\tau_1, \dots, \tau_r)\Big) = f\Big(\tilde{P}_q(\tau_1), \dots, \tilde{P}_q(\tau_r)\Big).$$

□

16.2.21. For $\sigma = (\sigma_1, \dots, \sigma_m)$ and a formal morphism $F : \mathbb{C}[V] \to \mathbb{C}[[X_1, \dots, X_p]]$ from $\mathbb{C}^p$ to V put $\sigma(F) := (F(\sigma_1), \dots, F(\sigma_m))$.

Lemma. *Let $F : \mathbb{C}[V] \to \mathbb{C}[[X_1, \dots, X_p]]$ be a formal morphism from $\mathbb{C}^p$ to V. If $\sigma(F) = 0$, then F vanishes on the set of all regular functions on V with zero constant terms; or, $F = 0$ as an element of $V \otimes \mathbb{C}[[X_1, \dots, X_p]]$.*

Proof. Let (e_i) be a basis of V and u_i the corresponding coordinates. It is sufficient to prove that $F(u_i) = 0$. Since the group G is finite the ring $\mathbb{C}[V]$ is integral over its subalgebra $\mathbb{C}[V]^G$. Then for each $1 \leq i \leq n$ there is a polynomial $p(x) = x^N + \sum_{j=1}^N a_{N-j} x^{N-j}$, whose coefficients a_{N-j} belong to $\mathbb{C}[V]^G$, such that $p(u_i) = 0$. Consider the natural grading of the ring $\mathbb{C}[V]$. Since $\deg((u_i)^N) = N$

we may assume that $\deg a_{N-j} = j$. This implies that the coefficients a_{N-j} as polynomials in the σ_j have no constant terms. Then we have

$$0 = F(p(u_i)) = F(u_i)^N + \sum_{j=1}^{n} F(a_{N-j})F(u_i)^{N-j}.$$

Since $\sigma(F) = 0$, this equation implies $F(u_i)^N = 0$ and therefore $F(u_i) = 0$. $\square$

16.2.22. The conditions for formal lifts. For $F \in J_0^\infty(\mathbb{C}^p, V)$ consider $P_q(\tau)(F)$ as a polynomial in $(t_A)_A$ with coefficients in $\mathbb{C}[[X_1, \dots, X_p]]$. Denote by $P_q(\tau)(F)_0(t)$ the polynomial in $(t_A)_A$ which is obtained by the evaluation of the coefficients of the polynomial $P_q(\tau)(F)$ at $X = (X_1, \dots, X_p) = 0$. Similarly, for $f \in J_0^\infty(\mathbb{C}^p, Z)$ consider $\tilde{P}_q(\tau)(F)$ as a polynomial in $(t_A)_A$ with coefficients in $\mathbb{C}((X_1, \dots, X_p))$ and denote by $\tilde{P}_q(\tau)(f)_0(t)$ the polynomial in t which is obtained by the evaluation of the coefficients of the polynomial $\tilde{P}_q(\tau)(f)$ at $X = (X_1, \dots, X_p) = 0$ whenever their values at $X = 0$ are defined.

For

$$A = (\underbrace{1, \dots, 1}_{r_1 \text{ times}}, \dots, \underbrace{p, \dots, p}_{r_p \text{ times}}),$$

put

$$t_A(X) := \frac{1}{r_1! \dots r_p!} \, (X_1)^{r_1} \dots (X_p)^{r_p}, \quad t(X) := (t_A(X))_{A \in \mathfrak{A}_{p,q}}.$$

For a formal power series $\varphi \in \mathbb{C}[[X_1, \dots, X_p]]$, denote by $\operatorname{Tay}^q \varphi$ the sum of the terms of ϕ of orders $\leq q$. Denote by $\tilde{\mathfrak{S}}_j(A_1, \dots, A_{d_j})$ the function $\tilde{\mathfrak{T}}(A_1, \dots, A_{d_j})$ for $\tau = \sigma_j$.

Recall that by, 16.2.6, for a quasi-regular formal morphism $f \in J_0^\infty(\mathbb{C}^p, Z)$ there is a choice of invariant coordinates such that for the corresponding function $\tilde{\Delta}$ we have $f(\tilde{\Delta}) \neq 0$ which here we write also as $\tilde{\Delta}(f) \neq 0$.

Theorem. *Let $f \in J_0^\infty(\mathbb{C}^p, Z)$ be a quasi-regular formal morphism given by the equations $f(w_j) = f_j \in \mathbb{C}[[X_1, \dots, X_p]]$ for $1 \leq j \leq m$, where w_j are the standard coordinate functions on $\mathbb{C}^m \supseteq Z$. Let y_i be invariant coordinates on V such that for the corresponding function $\tilde{\Delta}$ we have $\tilde{\Delta}(f) \neq 0$. Assume q is the minimal order of non-zero terms of $\tilde{\Delta}(f)$.*

Then a formal lift F of f exists if and only if for $1 \leq j \leq m$ and for each system of multi-indices $A_1, \dots, A_{d_j} \in \mathfrak{A}_{p,q}$ we have $\tilde{\mathfrak{S}}_j(A_1, \dots, A_{d_j})(f) \in \mathbb{C}[[X_1, \dots, X_p]]$ and $\operatorname{Tay}^q f_j = \tilde{P}_q(\sigma_j)(f)_0(t(X))$.

Proof. Let F be a formal lift of f. Then, by theorem 16.2.13, we have

$$\tilde{\mathfrak{S}}_j(A_1, \dots, A_{d_j})(f) = \mathfrak{S}_j(A_1, \dots, A_{d_j})(F) \in \mathbb{C}[[X_1, \dots, X_p]].$$

Moreover, by lemma 16.2.20, we have

$$\operatorname{Tay}^q f_j = \operatorname{Tay}^q(\sigma_j(F)) = P_q(\sigma_j)(F)_0(t(X)) = \tilde{P}_q(\sigma_j)(f)_0(t(X)).$$

Conversely, let the assumptions of the theorem be satisfied and let q be the minimal order of non-zero terms of $\tilde{\Delta}(f)$. Then $\operatorname{Tay}^q(\tilde{\Delta}(f)) \neq 0$. So there is a point $x_0 = (x_{0,1}, \dots, x_{0,p}) \in \mathbb{C}^p$ such that $\operatorname{Tay}^q(\tilde{\Delta}(f))(x_0) \neq 0$.

For each $1 \leq j \leq m$ consider the function

$$f_j^q((t_A)_A) := \tilde{P}_q(\sigma_j)(f) = \sum_{A_1, \dots, A_{d_j} \in \mathfrak{A}_{p,q}} \tilde{\mathfrak{S}}_j(A_1, \dots, A_{d_j})(f) \, t_{A_1} \dots t_{A_{d_j}}.$$

We may consider $f^q = (f_j^q)$ as a formal morphism $\mathbb{C}^{\mathfrak{A}_{p,q}} \times \mathbb{C}^p, (t(x_0)), 0) \to \mathbb{C}^m$, i.e., as a morphism $\tilde{\mathcal{O}}_{\mathbb{C}^m, f^q(t(x_0), 0)} \to \tilde{\mathcal{O}}_{\mathbb{C}^{\mathfrak{A}_{p,q}} \times \mathbb{C}^p, (t(x_0), 0)}$. We prove that $f^q = (f_j^q)$ is a formal morphism $\mathbb{C}^{\mathfrak{A}_{p,q}} \times \mathbb{C}^p, (t(x_0), 0) \to Z$ by the following arguments. Let

$\varPhi \in I(Z)$ be a homogeneous polynomial. Then $\varPhi \circ \sigma = 0$ and, by lemma 16.2.20, we have $\varPhi((f_j^q)) = \varPhi((\tilde{P}_q(\sigma_j)(f))) = \tilde{P}_q(\varPhi \circ (\sigma_j))(f) = 0$.

By assumption, we have $\mathrm{Tay}^q f_j = \tilde{P}_q(\sigma_j)(f)_0((t_A(X))_A) = f_j^q((t_A(X))_A, 0)$. Since the polynomial Δ is homogeneous, we have

$$\mathrm{Tay}^q\Big(\tilde{\Delta}(f^q(t(X),0))\Big)(x_0) = \mathrm{Tay}^q(\tilde{\Delta}(\mathrm{Tay}^q f))(x_0) = \mathrm{Tay}^q(\tilde{\Delta}(f))(x_0) \neq 0.$$

Thus, the formal morphism $f^q = (f_j^q)_{1\le j\le m}$ has a formal lift

$$F^q : \mathbb{C}^{\mathfrak{A}_{p,q}} \times \mathbb{C}^p, ((t(x_0), 0) \to V,$$

which can be written as $F^q = \sum_A F_A^q t_A$, where A is a multi-index and $F_A^q \in V \otimes \mathbb{C}[[X_1, \dots, X_p]]$, by proposition 16.2.14.

Since F^q is a formal lift of f^q, for each $1 \le j \le m$ we have

$$\sigma_j(F^q) = \sum_{A_1,\dots,A_{d_j}} \mathfrak{S}_j(F^q_{A_1}, \dots, F^q_{A_{d_j}})\, t_{A_1} \dots t_{A_{d_j}}$$

$$= \sum_{A_1,\dots,A_{d_j} \in \mathfrak{A}_{p,q}} \tilde{\mathfrak{S}}_j(A_1, \dots, A_{d_j})(f^q)\, (t_{A_1} + t_{A_1}(x_0)) \dots (t_{A_{d_j}} + t_{A_{d_j}}(x_0)).$$

This implies that $\mathfrak{S}_j(F^q_{A_1}, \dots, F^q_{A_{d_j}}) = 0$ whenever for some $1 \le k \le d_j$ we have $|A_k| > q$. In particular, for a multi-index A such that $|A| > q$ and for each $1 \le j \le m$ we have $\sigma_j(F_A^q) = 0$. By lemma 16.2.21, we have $F_A^q = 0$, and, therefore, the formal lift F^q is a polynomial in $(t_A)_A$ with coefficients in $V \otimes \mathbb{C}[[X_1, \dots, X_p]]$. Put $t_1 := (t_{1,A})_A$ where $t_{1,\emptyset} = 1$, and $t_{1,A} = 0$ for $A \neq \emptyset$. Denote by $F^q(t_1)$ the value of F^q as a polynomial in t at $t = t_1$. Then we have

$$\sigma_j(F^q(t_1)) = f_j^q(t_1) = \tilde{\mathfrak{S}}_j(\emptyset, \dots, \emptyset)(f) = f_j,$$

i.e., $F^q(t_1)$ is a formal lift of f. □

COROLLARY. *The map* $\pi^\infty : J_0^\infty(\mathbb{C}^p, V)/G \to J_0^\infty(\mathbb{C}^p, Z)$ *is injective.*

PROOF. Let $f \in J_0^\infty(\mathbb{C}^p, Z)$ be a formal morphism which has a lift to V.

First assume that the morphism f is quasi-regular. Consider a formal morphism $f^q = (f_j^q)_j$ from $\mathbb{C}^{\mathfrak{A}_{p,q}} \times \mathbb{C}^p, (t(x_0), 0)$ to Z constructed for f in the proof of theorem 16.2.22 and one of its lifts $F^q : \mathbb{C}^{\mathfrak{A}_{p,q}} \times \mathbb{C}^p, (t(x_0), 0) \to V$. Since $F^q(t(x_0), 0)$ is a regular point of V, the lift F^q is defined up to the action of some $g \in G$. On the other hand, for each lift F of f, $F^q = \sum_A \frac{1}{A!} \partial_A F\ t_A$ is a lift of f^q. This implies that the lift F of f is defined uniquely up to the action of some $g \in G$.

For an arbitrary formal morphism $f \in J_0^\infty(\mathbb{C}^p, Z)$, there is a subgroup K of G such that we can consider f as a quasi-regular formal morphism to $V^K/(N_G(K)/K)$. Then one can prove our statement using the same arguments as in the proof of theorem 16.2.6. □

16.2.23. In [**KLMR08**] the special case when G is a finite group generated by complex reflections is considered and treated in detail.

Bibliography

[AKLM98] D. Alekseevsky, A. Kriegl, M. Losik, and P. W. Michor, *Choosing roots of polynomials smoothly*, Israel J. Math. **105** (1998), 203–233.

[AKLM00] ———, *Lifting smooth curves over invariants for representations of compact Lie groups*, Transform. Groups **5** (2000), no. 2, 103–110.

[Arn86] V. I. Arnol′d, *Hyperbolic polynomials and Vandermonde mappings*, Funktsional. Anal. i Prilozhen. **20** (1986), no. 2, 52–53.

[AS81] M. Abud and G. Sartori, *The geometry of orbit-space and natural minima of Higgs potentials*, Phys. Lett. B **104** (1981), no. 2, 147–152.

[AS83] ———, *The geometry of spontaneous symmetry breaking*, Ann. Physics **150** (1983), no. 2, 307–372.

[Bal84] J. M. Ball, *Differentiability properties of symmetric and isotropic functions*, Duke Math. J. **51** (1984), no. 3, 699–728.

[Bar72] G. Barbançon, *Théorème de Newton pour les fonctions de class C^r*, Ann. Sci. École Norm. Sup. (4) **5** (1972), 435–457.

[Bar05] ———, *Whitney regularity of the image of the Chevalley mapping*, Preprint, `http://www.ma.utexas.edu/mp_arc/c/05/05-403.pdf`, 2005.

[Bau72] H. Baumgärtel, *Endlichdimensionale analytische Störungstheorie*, Akademie-Verlag, Berlin, 1972, Mathematische Lehrbücher und Monographien, II. Abteilung. Mathematische Monographien, Band 28.

[Bay04] T. Bayer, *Optimal descriptions of orbit spaces and strata of finite groups*, ISSAC 2004, ACM, New York, 2004, pp. 19–26.

[BBCP06] J.-M. Bony, F. Broglia, F. Colombini, and L. Pernazza, *Nonnegative functions as squares or sums of squares*, J. Funct. Anal. **232** (2006), no. 1, 137–147.

[BCR87] J. Bochnak, M. Coste, and M.-F. Roy, *Géométrie algébrique réelle*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 12, Springer-Verlag, Berlin, 1987.

[BGLS01] H. H. Bauschke, O. Güler, A. S. Lewis, and H. S. Sendov, *Hyperbolic polynomials and convex analysis*, Canad. J. Math. **53** (2001), no. 3, 470–488.

[Bie75] E. Bierstone, *Lifting isotopies from orbit spaces*, Topology **14** (1975), no. 3, 245–252.

[Bir71] D. Birkes, *Orbits of linear algebraic groups*, Ann. of Math. (2) **93** (1971), 459–475.

[BM00] E. Bierstone and P. D. Milman, *Subanalytic geometry*, Model theory, algebra, and geometry, Math. Sci. Res. Inst. Publ., vol. 39, Cambridge Univ. Press, Cambridge, 2000, pp. 151–172.

[Bom67] J. Boman, *Differentiability of a function and of its compositions with functions of one variable*, Math. Scand. **20** (1967), 249–268.

[Bon05] J.-M. Bony, *Sommes de carrés de fonctions dérivables*, Bull. Soc. Math. France **133** (2005), no. 4, 619–639.

[Bor66] A. Borel, *Linear algebraic groups*, Algebraic Groups and Discontinuous Subgroups (Proc. Sympos. Pure Math., Boulder, Colo., 1965), Amer. Math. Soc., Providence, R.I., 1966, pp. 3–19.

[Bor69] ———, *Linear algebraic groups*, W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[BR83] G. Barbançon and M. Raïs, *Sur le théorème de Hilbert différentiable pour les groupes linéaires finis (d'après E. Noether)*, Ann. Sci. École Norm. Sup. (4) **16** (1983), no. 3, 355–373 (1984).

[Bre72] G. E. Bredon, *Introduction to compact transformation groups*, Academic Press, New York, 1972, Pure and Applied Mathematics, Vol. 46.

[Bre93] ———, *Topology and geometry*, Graduate Texts in Mathematics, vol. 139, Springer-Verlag, New York, 1993.

[Bro79] M. D. Bronshtein, *Smoothness of roots of polynomials depending on parameters*, Sibirsk. Mat. Zh. **20** (1979), no. 3, 493–501, 690, English transl. in Siberian Math. J. **20** (1980), 347–352. See also MR0537355 (82c:26018).

[Bro80] ———, *The Cauchy problem for hyperbolic operators with characteristics of variable multiplicity*, Trudy Moskov. Mat. Obshch. **41** (1980), 83–99, English transl. in Trans. Moscow Math. Soc. **1**(1982), 87–103.

[Bru79] G. W. Brumfiel, *Partially ordered rings and semi-algebraic geometry*, London Mathematical Society Lecture Note Series, vol. 37, Cambridge University Press, Cambridge, 1979.

[BS73a] J.-M. Bony and P. Schapira, *Existence et prolongement des solutions holomorphes des équations aux dérivées partielles*, Colloque International CNRS sur les Équations aux Dérivées Partielles Linéaires (Univ. Paris-Sud, Orsay, 1972), Soc. Math. France, Paris, 1973, pp. 117–127. Astérisque, 2 et 3.

[BS73b] ———, *Solutions analytiques et hyperfonctions du problème de Cauchy pour les opérateurs hyperboliques non stricts*, Colloque International CNRS sur les Équations aux Dérivées Partielles Linéaires (Univ. Paris-Sud, Orsay, 1972), Soc. Math. France, Paris, 1973, pp. 108–116. Astérisque, 2 et 3.

[CC04] J. Chaumat and A.-M. Chollet, *Division par un polynôme hyperbolique*, Canad. J. Math. **56** (2004), no. 6, 1121–1144.

[Che46] C. Chevalley, *Theory of Lie Groups. I*, Princeton Mathematical Series, vol. 8, Princeton University Press, Princeton, N. J., 1946.

[Che55] ———, *Théorie des groupes de Lie. Tome III. Théorèmes généraux sur les algèbres de Lie*, Actualités Sci. Ind. no. 1226, Hermann & Cie, Paris, 1955.

[ĆM06] B. Ćurgus and V. Mascioni, *Roots and polynomials as homeomorphic spaces*, Expo. Math. **24** (2006), no. 1, 81–95.

[CP88] C. J. Cummins and J. Patera, *Polynomial icosahedral invariants*, J. Math. Phys. **29** (1988), no. 8, 1736–1745.

[Dad85] J. Dadok, *Polar coordinates induced by actions of compact Lie groups*, Trans. Amer. Math. Soc. **288** (1985), no. 1, 125–137.

[Die60] J. Dieudonné, *Foundations of modern analysis*, Pure and Applied Mathematics, Vol. X, Academic Press, New York, 1960.

[Die70] ———, *Sur un théorème de Glaeser*, J. Analyse Math. **23** (1970), 85–88.

[DK85] J. Dadok and V. Kac, *Polar representations*, J. Algebra **92** (1985), no. 2, 504–524.

[DK02] H. Derksen and G. Kemper, *Computational invariant theory*, Invariant Theory and Algebraic Transformation Groups, I, Springer-Verlag, Berlin, 2002, Encyclopaedia of Mathematical Sciences, 130.

[FF89] C.-A. Faure and A. Frölicher, *Hölder differentiable maps and their function spaces*, Categorical topology and its relation to analysis, algebra and combinatorics (Prague, 1988), World Sci. Publ., Teaneck, NJ, 1989, pp. 135–142.

[FK88] A. Frölicher and A. Kriegl, *Linear spaces and differentiation theory*, Pure and Applied Mathematics (New York), John Wiley & Sons Ltd., Chichester, 1988, A Wiley-Interscience Publication.

[FP78] C. Fefferman and D. H. Phong, *On positivity of pseudo-differential operators*, Proc. Nat. Acad. Sci. U.S.A. **75** (1978), no. 10, 4673–4674.

[Frö80] A. Frölicher, *Catégories cartésiennement fermées engendrées par des monoïdes*, Cahiers Topologie Géom. Différentielle **21** (1980), no. 4, 367–375, Third Colloquium on Categories (Amiens, 1980), Part I.

[Frö81] ———, *Applications lisses entre espaces et variétés de Fréchet*, C. R. Acad. Sci. Paris Sér. I Math. **293** (1981), no. 2, 125–127.

[Går51] L. Gårding, *Linear hyperbolic partial differential equations with constant coefficients*, Acta Math. **85** (1951), 1–62.

[Går59] ———, *An inequality for hyperbolic polynomials*, J. Math. Mech. **8** (1959), 957–965.

[GB85] L. C. Grove and C. T. Benson, *Finite reflection groups*, second ed., Graduate Texts in Mathematics, vol. 99, Springer-Verlag, New York, 1985.

[Giv87] A. B. Givental′, *Moments of random variables and the equivariant Morse lemma*, Uspekhi Mat. Nauk **42** (1987), no. 2(254), 221–222.

[Gla63a] G. Glaeser, *Fonctions composées différentiables*, Ann. of Math. (2) **77** (1963), 193–209.

[Gla63b] ———, *Racine carrée d'une fonction différentiable*, Ann. Inst. Fourier (Grenoble) **13** (1963), no. fasc. 2, 203–210.

[Hil88] D. Hilbert, *über die Darstellung definiter Formen als Summe von Formenquadraten*, Math. Ann. **32** (1888), no. 3, 342–350, see also Gesammelte Abhandlungen, Bd. 2 (Chelsea Publishing Company, Bronx, New York 1965), 154–161.

[HLP52] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, second ed., Cambridge University Press, Cambridge, 1952.

[HM59] G. Hochschild and G. D. Mostow, *Representations and representative functions of Lie groups. III*, Ann. of Math. (2) **70** (1959), 85–100.

[Hoc65] G. Hochschild, *The structure of Lie groups*, Holden-Day Inc., San Francisco, 1965.

[Hör63] L. Hörmander, *Linear partial differential operators*, Die Grundlehren der mathematischen Wissenschaften, Bd. 116, Academic Press Inc., Publishers, New York, 1963.

[Hör83a] ———, *The analysis of linear partial differential operators. I*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 256, Springer-Verlag, Berlin, 1983, Distribution theory and Fourier analysis.

[Hör83b] ———, *The analysis of linear partial differential operators. II*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 257, Springer-Verlag, Berlin, 1983, Differential operators with constant coefficients.

[Hör94] ———, *Notions of convexity*, Progress in Mathematics, vol. 127, Birkhäuser Boston Inc., Boston, MA, 1994.

[Hum90] J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Studies in Advanced Mathematics, vol. 29, Cambridge University Press, Cambridge, 1990.

[Kat76] T. Kato, *Perturbation theory for linear operators*, second ed., Grundlehren der Mathematischen Wissenschaften, vol. 132, Springer-Verlag, Berlin, 1976.

[Kem78] G. R. Kempf, *Instability in invariant theory*, Ann. of Math. (2) **108** (1978), no. 2, 299–316.

[KLM03] A. Kriegl, M. Losik, and P. W. Michor, *Tensor fields and connections on holomorphic orbit spaces of finite groups*, J. Lie Theory **13** (2003), no. 2, 519–534.

[KLM04] ———, *Choosing roots of polynomials smoothly. II*, Israel J. Math. **139** (2004), 183–188.

[KLMR05] A. Kriegl, M. Losik, P. W. Michor, and A. Rainer, *Lifting smooth curves over invariants for representations of compact Lie groups. II*, J. Lie Theory **15** (2005), no. 1, 227–234.

[KLMR06] ———, *Lifting smooth curves over invariants for representations of compact Lie groups. III*, J. Lie Theory **16** (2006), no. 3, 579–600.

[KLMR08] ———, *Lifting mappings over invariants of finite groups*, Acta Math. Univ. Comenian. (N.S.) **77** (2008), no. 1, 93–122.

[KM97] A. Kriegl and P. W. Michor, *The convenient setting of global analysis*, Mathematical Surveys and Monographs, vol. 53, American Mathematical Society, Providence, RI, 1997, `http://www.ams.org/online_bks/surv53/`.

[KM03] ———, *Differentiable perturbation of unbounded operators*, Math. Ann. **327** (2003), no. 1, 191–201.

[KMS93] I. Kolář, P. W. Michor, and J. Slovák, *Natural operations in differential geometry*, Springer-Verlag, Berlin, 1993.

[KN79] G. Kempf and L. Ness, *The length of vectors in representation spaces*, Algebraic geometry (Proc. Summer Meeting, Univ. Copenhagen, Copenhagen, 1978), Lecture Notes in Math., vol. 732, Springer, Berlin, 1979, pp. 233–243.

[Kna96] A. W. Knapp, *Lie groups beyond an introduction*, Progress in Mathematics, vol. 140, Birkhäuser Boston Inc., Boston, MA, 1996.

[Kom77] H. Komatsu, *Irregularity of characteristic elements and hyperbolicity*, Proceedings of the Oji Seminar on Algebraic Analysis and the RIMS Symposium on Algebraic Analysis (Kyoto Univ., Kyoto, 1976), vol. 12, 1976/77, pp. 233–245.

[Kos53] J. L. Koszul, *Sur certains groupes de transformations de Lie*, Géométrie différentielle. Colloques Internationaux du Centre National de la Recherche Scientifique, Strasbourg, 1953, Centre National de la Recherche Scientifique, Paris, 1953, pp. 137–141.

[Kos89] V. P. Kostov, *On the geometric properties of Vandermonde's mapping and on the problem of moments*, Proc. Roy. Soc. Edinburgh Sect. A **112** (1989), no. 3-4, 203–211.

[KP08] K. Kurdyka and L. Paunescu, *Hyperbolic polynomials and multiparameter real-analytic perturbation theory*, Duke Math. J. **141** (2008), no. 1, 123–149.

[Lan02] S. Lang, *Algebra*, third ed., Graduate Texts in Mathematics, vol. 211, Springer-Verlag, New York, 2002.

[Lew96] A. S. Lewis, *Convex analysis on the Hermitian matrices*, SIAM J. Optim. **6** (1996), no. 1, 164–177.

[LMP03] M. Losik, P. W. Michor, and V. L. Popov, *Invariant tensor fields and orbit varieties for finite algebraic transformation groups*, A tribute to C. S. Seshadri (Chennai, 2002), Trends Math., Birkhäuser, Basel, 2003, pp. 346–378.

[Loj65] S. Lojasiewicz, *Ensembles semi-analytiques*, Inst. Hautes Études Sci., Bures-sur-Yvette, 1965.

[Los01] M. Losik, *Lifts of diffeomorphisms of orbit spaces for representations of compact Lie groups*, Geom. Dedicata **88** (2001), no. 1-3, 21–36.

[LR07] M. Losik and A. Rainer, *Choosing roots of polynomials with symmetries smoothly*, Rev. Mat. Complut. **20** (2007), no. 2, 267–291.

[Lun73] D. Luna, *Slices étales*, Sur les groupes algébriques, Soc. Math. France, Paris, 1973, pp. 81–105. Bull. Soc. Math. France, Paris, Mémoire 33.

[Lun76] ———, *Fonctions différentiables invariantes sous l'opération d'un groupe réductif*, Ann. Inst. Fourier (Grenoble) **26** (1976), no. 1, ix, 33–49.

[Lya83] O. V. Lyashko, *Geometry of bifurcation diagrams*, Current problems in mathematics, Vol. 22, Itogi Nauki i Tekhniki, Akad. Nauk SSSR Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1983, pp. 94–129.

[Man85] T. Mandai, *Smoothness of roots of hyperbolic polynomials with respect to one-dimensional parameter*, Bull. Fac. Gen. Ed. Gifu Univ. (1985), no. 21, 115–118.

[Mat60] Y. Matsushima, *Espaces homogènes de Stein des groupes de Lie complexes*, Nagoya Math. J **16** (1960), 205–218.

[Mat73] J. N. Mather, *Stratifications and mappings*, Dynamical systems (Proc. Sympos., Univ. Bahia, Salvador, 1971), Academic Press, New York, 1973, pp. 195–232.

[Mat77] ———, *Differentiable invariants*, Topology **16** (1977), no. 2, 145–155.

[Meg92] I. Meguerditchian, *A theorem on the escape from the space of hyperbolic polynomials*, Math. Z. **211** (1992), no. 3, 449–460.

[Meh88] M. L. Mehta, *Basic sets of invariant polynomials for finite reflection groups*, Comm. Algebra **16** (1988), no. 5, 1083–1098.

[Mic97] P. W. Michor, *Isometric actions of lie groups and invariants*, Lecture notes of a course held in Vienna, 1993, 1997, `http://www.mat.univie.ac.at/~michor/tgbook.pdf`, 1997.

[Mos55] G. D. Mostow, *Self-adjoint groups*, Ann. of Math. (2) **62** (1955), 44–55.

[Mum88] D. Mumford, *The red book of varieties and schemes*, Lecture Notes in Mathematics, vol. 1358, Springer-Verlag, Berlin, 1988.

[MY57] D. Montgomery and C. T. Yang, *The existence of a slice*, Ann. of Math. (2) **65** (1957), 108–116.

[Noe15] E. Noether, *Der endlichkeitssatz der invarianten endlicher gruppen*, Math. Ann. **77** (1915), 89–92 (German).

[Nui68] W. Nuij, *A note on hyperbolic polynomials*, Math. Scand. **23** (1968), 69–72 (1969).

[Ost40] A. Ostrowski, *Recherches sur la méthode de Graeffe et les zéros des polynomes et des séries de Laurent*, Acta Math. **72** (1940), 99–155.

[Pal61] R. S. Palais, *On the existence of slices for actions of non-compact Lie groups*, Ann. of Math. (2) **73** (1961), 295–323.

[Pfl01] M. J. Pflaum, *Analytic and geometric study of stratified spaces*, Lecture Notes in Mathematics, vol. 1768, Springer-Verlag, Berlin, 2001.

[Pro78] C. Procesi, *Positive symmetric functions*, Adv. in Math. **29** (1978), no. 2, 219–225.

[PS85] C. Procesi and G. Schwarz, *Inequalities defining orbit spaces*, Invent. Math. **81** (1985), no. 3, 539–554.

[PSW78] J. Patera, R. T. Sharp, and P. Winternitz, *Polynomial irreducible tensors for point groups*, J. Math. Phys. **19** (1978), no. 11, 2362–2376.

[PT87] R. S. Palais and C.-L. Terng, *A general theory of canonical forms*, Trans. Amer. Math. Soc. **300** (1987), no. 2, 771–789.

[PT88] ———, *Critical point theory and submanifold geometry*, Lecture Notes in Mathematics, vol. 1353, Springer-Verlag, Berlin, 1988.

[Rel37] F. Rellich, *Störungstheorie der Spektralzerlegung*, Math. Ann. **113** (1937), no. 1, 600–619.

[Rel42] ———, *Störungstheorie der Spektralzerlegung. V*, Math. Ann. **118** (1942), 462–484.

[Rel69] ———, *Perturbation theory of eigenvalue problems*, Assisted by J. Berkowitz. With a preface by Jacob T. Schwartz, Gordon and Breach Science Publishers, New York, 1969.

[Riv74] T. J. Rivlin, *The Chebyshev polynomials*, Wiley-Interscience [John Wiley & Sons], New York, 1974, Pure and Applied Mathematics.

[Roc70] R. T. Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970.

[RS02] Q. I. Rahman and G. Schmeisser, *Analytic theory of polynomials*, London Mathematical Society Monographs. New Series, vol. 26, The Clarendon Press Oxford University Press, Oxford, 2002.

[Rud87] W. Rudin, *Real and complex analysis*, third ed., McGraw-Hill Book Co., New York, 1987.

[Sar83] G. Sartori, *A theorem on orbit structures (strata) of compact linear Lie groups*, J. Math. Phys. **24** (1983), no. 4, 765–768.

[Sch75] G. W. Schwarz, *Smooth functions invariant under the action of a compact Lie group*, Topology **14** (1975), 63–68.

[Sch80] ———, *Lifting smooth homotopies of orbit spaces*, Inst. Hautes Études Sci. Publ. Math. (1980), no. 51, 37–135.

[Smi95] L. Smith, *Polynomial invariants of finite groups*, Research Notes in Mathematics, vol. 6, A K Peters Ltd., Wellesley, MA, 1995.

[Spa99] S. Spagnolo, *On the absolute continuity of the roots of some algebraic equations*, Ann. Univ. Ferrara Sez. VII (N.S.) **45** (1999), no. suppl., 327–337 (2000), Workshop on Partial Differential Equations (Ferrara, 1999).

[SS87] L. Smith and R. E. Stong, *On the invariant theory of finite groups: orbit polynomials and splitting principles*, J. Algebra **110** (1987), no. 1, 134–157.

[Sta82] J. Stasica, *The Whitney condition for subanalytic sets*, Zeszyty Nauk. Uniw. Jagielloń. Prace Mat. (1982), no. 23, 211–221.

[Stu93] B. Sturmfels, *Algorithms in invariant theory*, Texts and Monographs in Symbolic Computation, Springer-Verlag, Vienna, 1993.

[SV03] G. Sartori and G. Valente, *Tools in the orbit space approach to the study of invariant functions: rational parametrization of strata*, J. Phys. A **36** (2003), no. 7, 1913–1929.

[Ter85] C.-L. Terng, *Isoparametric submanifolds and their Coxeter groups*, J. Differential Geom. **21** (1985), no. 1, 79–107.

[Tou72] J.-C. Tougeron, *Idéaux de fonctions différentiables*, Springer-Verlag, Berlin, 1972, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 71.

[Ver83] P. Ver Eecke, *Fondements du calcul différentiel*, Mathématiques. [Mathematics], Presses Universitaires de France, Paris, 1983.

[VP89] È. B. Vinberg and V. L. Popov, *Invariant theory*, Algebraic geometry, 4 (Russian), Itogi Nauki i Tekhniki, Akad. Nauk SSSR Vsesoyuz. Inst. Nauchn. i Tekhn. Inform., Moscow, 1989, pp. 137–314, 315.

[Wak86] S. Wakabayashi, *Remarks on hyperbolic polynomials*, Tsukuba J. Math. **10** (1986), no. 1, 17–28.

[Wey39] H. Weyl, *The Classical Groups. Their Invariants and Representations*, Princeton University Press, Princeton, N.J., 1939.

[Whi34] H. Whitney, *Functions differentiable on the boundaries of regions*, Ann. of Math. (2) **35** (1934), no. 3, 482–485.

[Whi43] ———, *Differentiable even functions*, Duke Math. J. **10** (1943), 159–160.

Index