Choosing roots of polynomials smoothly and lifting smooth curves over invariants

Dissertation
zur Erlangung des Akademischen Grades
'Doktor der Naturwissenschaften'
an der Naturwissenschaftlichen Fakultät
der Universität Wien

eingereicht von
Armin Rainer

betreut von
Ao.Univ.-Prof. Dr. Peter W. Michor

Wien, am 27. Januar 2004

Einleitung

Der Ausgangspunkt und die zugrundeliegende Fragestellung dieser Arbeit ist ein Problem, das eine gewisse Rolle in der Theorie der Partiellen Differentialgleichungen in Verbindung mit dem sogenannten Cauchy-Problem spielt. Aber dieses Problem ist auch für sich allein interessant. Man betrachte eine Kurve von Polynomen mit festem Grad n

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

mit ausschließlich reellen Wurzeln (später werden wir solche Polynome 'hyperbolisch' nennen), welche durch t nahe 0 in $\mathbb{R}$ glatt parametrisiert ist. Können wir n glatte Funktionen $x_1(t), \dots, x_n(t)$ finden, die die Wurzeln von $P(t)$ für jedes t parametrisieren?

Dem Studium dieses Problems ist der erste Teil der vorliegenden Arbeit gewidmet. Es handelt sich dabei vorwiegend um analytische und funktionentheoretische Überlegungen. Nur im Kapitel 4 streifen wir kurz die Theorie der Partiellen Differentialgleichungen.

Im Kapitel 1 werden einige Stetigkeitsresultate der Wurzeln von Polynomen in Abhängigkeit ihrer Koeffizienten vorgestellt. Sie werden dann vorwiegend in den Kapiteln 3 und 4 verwendet. Ein weitreichenderes Stetigkeitsresultat, welches eine stetige Parametrisierung der Wurzeln von hyperbolischen Polynomen liefert, wird schließlich im Abschnitt 2.4 bewiesen. Es gehört zu einer gut strukturierten Behandlung des Problems, die auf Alekseevsky, Kriegl, Losik und Michor [**1**] (1998) zurückgeht und den Inhalt des Kapitels 2 ausmacht. Dieses Kapitel enthält viele interessante Resultate wie die Beschreibung des Raumes der hyperbolischen Polynome eines festen Grades und die Lösung unseres Problems unter recht allgemeinen Bedingungen.

Nichtsdestotrotz waren die oben genannten nicht in der Lage, mit ihren Methoden ein Resultat zu zeigen, das schon 1979 von Bronshtein [**8**] gefunden und 1986 von Wakabayashi [**41**] in einfacherer Weise neu bewiesen worden war. Das erwähnte Resultat sagt aus, daß die Wurzeln einer C^n-Kurve von hyperbolischen Polynomen vom Grad n differenzierbar mit lokal beschränkter Ableitung gewählt werden können. Weil Bronshteins Diskussion des Problems, welche im Kapitel 3 vorgestellt wird, ziemlich lang, verwickelt und technisch ist, behandeln wir im Abschnitt 3.1 den Fall $n = 3$ sehr ausführlich, aber unter Verwendung der gesamten Argumentationsmaschinerie. Das Kapitel 4 ist den Methoden Wakabayashis gewidmet. Die Notation und der Hintergrund dieses Kapitels basieren auf Hörmander [**14**] und [**15**].

Der erste Teil dieser Abhandlung endet mit Kapitel 5, in welchem das Resultat von Bronshtein bzw. Wakabayashi verwendet wird, um zu zeigen, daß jede differenzierbare Parametrisierung der Wurzeln einer C^{2n}-Kurve von hyperbolischen Polynomen vom Grad n eigentlich schon C^1 ist und daß es immer eine zweimal differenzierbare Parametrisierung der Wurzeln einer C^{3n}-Kurve von hyperbolischen Polynomen vom Grad n gibt. Es ist bemerkenswert, daß diese Konklusionen bestmöglich sind. Diese Ergebnisse stammen von Kriegl, Losik und Michor [**17**] (2002).

Im zweiten Teil der vorliegenden Arbeit behandeln wir eine Verallgemeinerung des obigen Problems, die durch folgende Sichtweise motiviert ist: Die symmetrische Gruppe S_n wirke auf dem $\mathbb{R}^n$ durch Permutation der Koordinaten (die Wurzeln von P). Man betrachte die polynomiale Abbildung $\sigma = (\sigma_1, \ldots, \sigma_n) : \mathbb{R}^n \to \mathbb{R}^n$, deren Komponenten die elementar-symmetrischen Polynome (die Koeffizienten von P) sind. Unter diesem Blickwinkel lautet unsere Fragestellung: Gegeben eine glatte Kurve $c : \mathbb{R} \to \sigma(\mathbb{R}^n) \subseteq \mathbb{R}^n$, ist es möglich, einen glatten Lift $\bar{c} : \mathbb{R} \to \mathbb{R}^n$ von c zu finden, d.h. eine glatte Kurve $\bar{c}$, die $\sigma \circ \bar{c} = c$ erfüllt?

Im allgemeinen betrachten wir nun eine orthogonale Darstellung einer kompakten Liegruppe G auf einem reellen endlichdimensionalen Euklidischen Vektorraum V. Sei $\sigma_1, \ldots, \sigma_n$ ein System von homogenen Erzeugern der Algebra $\mathbb{R}[V]^G$ der invarianten Polynome auf V. Dann induziert die Abbildung $\sigma = (\sigma_1, \ldots, \sigma_n) : V \to \mathbb{R}^n$ eine Identifizierung des Orbitraumes V/G mit der semialgebraischen Menge $\sigma(V) \subseteq \mathbb{R}^n$. Nun können wir fragen: Gegeben eine glatte Kurve $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ im Orbitraum (glatt als Kurve im $\mathbb{R}^n$), gibt es einen glatten Lift nach V, d.h. eine glatte Kurve $\bar{c}$, die $\sigma \circ \bar{c} = c$ erfüllt?

Im Kapitel 6 wird der Hintergrund aus der Theorie der isometrischen Wirkungen vom Liegruppen bereitgestellt, welcher in den folgenden Kapiteln gebraucht wird. Dieses Kapitel beinhaltet eine Charakterisierung des Orbitraumes $V/G = \sigma(V)$, den differenzierbaren Scheibensatz, eine detailierte Behandlung der Isotropiedarstellung, sowie die Stratifizierung des Orbitraumes. In diesem Kapitel habe ich versucht eine möglichst allgemeine Darstellung der Theorie zu liefern, ohne jedoch zu weit von unserer Fragestellung, dem Liftungsproblem, abzuschweifen.

Mehrere Ergebnisse zu diesem Liftungsproblem werden im Kapitel 7 präsentiert, das sich auf eine Arbeit von Alekseevsky, Kriegl, Losik und Michor [**2**] aus dem Jahre 2002 stützt. Die wichtigsten darunter sind: eine reell analytische Kurve in V/G erlaubt einen lokalen reell analytischen Lift nach V; eine glatte Kurve in V/G erlaubt einen globalen glatten Lift, wenn gewisse Generizitätsbedingungen, welche die Statifizierung des Orbitraumes involvieren, erfüllt sind, siehe Abschnitt 7.1; und in beiden Fällen können die Lifts global, orthogonal zu jedem Orbit, das sie treffen, und eindeutig bis auf eine Transformation von G gewählt werden, wenn die Darstellung von G auf V polar ist, siehe Abschnitt 7.3.

Die Analyse des Liftungsproblems wird im Kapitel 8 fortgesetzt, wobei nun der Kurve im Orbitraum schwächere Differenzierbarkeitseigenschaften auferlegt werden. In erster Linie werden die erwähnten Generizitätsbedingungen weggelassen. Wir zeigen, daß eine stetige Kurve im Orbitraum stetig nach V geliftet werden kann und daß eine hinreichend oft differenzierbare Kurve in V/G einen globalen einmal differenzierbaren Lift nach V zuläßt. Was wir mit 'hinreichend oft differenzierbar' meinen, wird im Abschnitt 8.3 erklärt. Darüber hinaus liefert der Abschnitt 8.4 sogar einen orthogonalen differenzierbaren Lift einer glatten Kurve im Orbitraum. Das Kapitel 8 basiert auf einer Arbeit von Kriegl, Losik, Michor und Rainer [**19**], die bald erscheinen wird.

Dieses letzte Kapitel endet mit einem Ausblick (Abschnitt 8.5), in dem offene Fragen angesprochen werden. Das anspruchsvollste offene Problem and auch das Fernziel ist, die Existenz eines zweimal differenzierbaren Lifts im allgemeinen Setting ohne die erwähnten Generizitätsbedingungen zu zeigen. Mehr können wir nicht erwarten. Der Schlüssel dazu ist die Verallgemeinerung des Resultats von Bronshtein bzw. Wakabayashi aus den Kapiteln 3 und 4.

Mit wenigen Ausnahmen ist die Abhandlung in sich selbst geschlossen. Jene Resultate, welche ohne Beweis präsentiert werden, sind mit Verweisen an das Literaturverzeichnis versehen, und meistens sind sie für die Entwicklung und die Zielsetzung der Arbeit nur am Rande bedeutsam.

Mein Dank gilt im Besonderen meinem Betreuer Peter W. Michor für hilfreiche Anregung und Unterstützung. Weiters danken möchte ich Andreas Cap, Stefan Haller, Andreas Kriegl und Mark Losik für die zahlreichen fruchtbaren Diskussionen.

Preface

The basic and starting problem of this treatise is a question which plays a certain role in the theory of partial differential equations in connection with the so-called Cauchy problem, but is also interesting on its own. Consider a curve of polynomials of fixed degree n

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

with only real roots (later on we will say 'hyperbolic') and smoothly parameterized by t near 0 in $\mathbb{R}$. Can we find n smooth functions $x_1(t), \ldots, x_n(t)$ which parameterize the roots of $P(t)$ for each t?

To the study of this problem is dedicated the first part of the present work. It is mostly elementary calculus and complex analysis and does only slightly touch the theory of partial differential equations in chapter 4.

In chapter 1 a few continuity results of the roots of polynomials depending on their coefficients are presented. They will be used mainly in chapter 3 and in chapter 4. A further reaching continuity result is proved in section 2.4, providing a continuous parameterization of the roots of hyperbolic polynomials. It makes part of a well structured approach to the problem which is due to Alekseevsky, Kriegl, Losik and Michor [**1**] (1998) and is the content of chapter 2. It includes many interesting results as the description of the space of hyperbolic polynomials of a fixed degree and the solution of the problem under quite general conditions.

Nevertheless they where not able to show with their methods a result already found in 1979 by Bronshtein [**8**] and proved again in a different easier way in 1986 by Wakabayashi [**41**]. The mentioned result states that the roots of a C^n-curve of hyperbolic polynomials of degree n may be chosen differentiable with locally bounded derivative. Since Bronshtein's approach, presented in chapter 3, is quite long, involved and technical, we discuss in section 3.1 the case $n = 3$ at great length using the whole machinery of his argumentation. Chapter 4 is devoted to Wakabayashi's approach. Notation and background in this chapter are based on Hörmander [**14**] and [**15**].

Part 1 of this treatise is concluded by chapter 5 in which Bronshtein's and Wakabayashi's result is used to prove that any differentiable parameterization of the roots of a C^{2n}-curve of hyperbolic polynomials of degree n is actually C^1, and that there is always a twice differentiable parameterization of the roots of a C^{3n}-curve of hyperbolic polynomials of degree n. It is remarkable that these conclusions are best possible. These results are due to Kriegl, Losik and Michor [**17**] (2002).

In part 2 we treat a generalization of the above problem which is motivated by the following point of view (see section 6.1): Let the symmetric group S_n act in $\mathbb{R}^n$ by permuting the coordinates (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 (the coefficients of P). Now the question of interest reformulates to: 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 $\sigma \circ \bar{c} = c$?

We consider now in general 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 $\sigma(V) \subseteq \mathbb{R}^n$. Now we may ask: Given a smooth curve $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ in the orbit space (smooth as curve in $\mathbb{R}^n$), does there exist a smooth lift to V, i.e., a smooth curve $\bar{c} : \mathbb{R} \to V$ satisfying $\sigma \circ \bar{c} = c$?

Chapter 6 sets out the background from the theory of isometric actions of Lie groups required in the later chapters. It includes a characterization of the orbit space $V/G = \sigma(V)$, the differentiable slice theorem, a detailed treatment of the isotropy representation, and the stratification of the orbit space. In this chapter I tried to give an as general as possible description of the theory and at the same time not to depart too much from our subject, the lifting problem.

Many results concerning this lifting problem are presented in chapter 7 which is based on a paper of Alekseevsky, Kriegl, Losik and Michor [**2**] published in 2002. The most important are: a real analytic curve in V/G admits a local real analytic lift to V, a smooth curve in V/G admits a global smooth lift, if certain genericity conditions involving the stratification of the orbit space are satisfied, see section 7.1, and in both cases the lifts may be chosen global, orthogonal to each orbit they meet, and unique up to a transformation from G, whenever the representation of G on V is polar, see section 7.3.

The analysis of the lifting problem is continued in chapter 8, where weaker differentiability conditions are imposed on the curve in the orbit space. Primarily, the mentioned genericity conditions are omitted. It is shown that a continuous curve in the orbit space can be lifted to V continuously, and that a sufficiently often differentiable curve in V/G allows a global once differentiable lift to V. What is meant by 'sufficiently often differentiable' is explained in section 8.3. Moreover, section 8.4 provides even an orthogonal differentiable lift of a smooth curve in the orbit space. Chapter 8 is based on a paper of Kriegl, Losik, Michor and Rainer [**19**] which will be published soon.

This last chapter is concluded with an outlook (section 8.5) containing open problems. The most challenging open problem and the long-term object is to prove the existence of a twice differentiable lift in the general setting without the mentioned genericity conditions. We cannot expect more. The key is the generalization of Bronshtein's and Wakabayashi's result from chapter 3 and chapter 4.

With only a few exceptions the treatise is self-contained. 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.

Special thanks go to my supervisor Peter W. Michor for helpful encouragement. Moreover, I would like to thank Andreas Cap, Stefan Haller, Andreas Kriegl and Mark Losik for many fruitful discussions.

Contents

Part 1

Choosing roots of polynomials smoothly

CHAPTER 1

Continuity of the roots

1.1. A first continuity theorem

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 $\mathbb{C}$. A first approach to this problem is nearly trivial:

PROPOSITION 1.1.1. *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 $i = 0, \ldots, n-1$, 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 another monic polynomial ($b_n = 1$) with roots $z_1, \ldots, z_n$. 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

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

So indeed 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 $i = 0, \ldots, n-1$, have a root z satisfying $|w - z| < \epsilon$. □

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 1.1.2. *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

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

Arranging the roots of P as $|z_1| \le \cdots \le |z_n|$, 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 $z_{m+1}, \ldots, z_n$. 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

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

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

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

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

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

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

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

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

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

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

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 $i = 0, \dots, m-1$, we have $|a_i| = |a_i|^{\frac{n-i}{n-i}} \leq \max_{0 \leq j \leq m-1} |a_j|^{\frac{n-i}{n-j}}$ which implies

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

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

Now we are prepared to show the following deeper theorem on the continuity of roots of polynomials as functions of the coefficients.

THEOREM 1.1.3. *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 $x_1, \dots, x_p$ of multiplicities $m_1, \dots, m_p$. Then, given a positive $\epsilon < \min_{1 \leq i < j \leq p} \frac{|x_i - x_j|}{2}$, 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 $i = 0, \dots, n-1$, has exactly m_j roots in the disk

$$D(x_j; \epsilon) = \{z \in \mathbb{C} : |z - x_j| \leq \epsilon\} \qquad (j = 1, \dots, p).$$

More precisely: Let

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

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,

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

Now, applying lemma 1.1.2 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 lemma 1.1.2, the moduli of the roots of P are bounded by A. Since $A \ge 1$ and $\binom{k}{i} < \sum_{l=0}^{k} \binom{k}{l} = 2^k$ for $i = 0, \ldots, k$, 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 theorem 1.1.3, we may say that the roots of a polynomial of degree n, as functions of the coefficients, satisfy a Lipschitz condition of order $\frac{1}{n}$.

1.2. Rouché's theorem and an application

Another possibility to get results on the continuity of roots 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.

Suppose a function $f : \mathbb{C} \to \mathbb{C}$ is holomorphic in $a \in \mathbb{C}$ or has an isolated singularity in a. Then the *logarithmic residue of f in a* is defined to be the residue of the logarithmic derivative $(\log \circ f)' = \frac{f'}{f}$ in a.
For an n-fold root a of f and z near a we have

$$f(z) = c_n(z-a)^n + c_{n+1}(z-a)^{n+1} + \cdots \qquad (c_n \neq 0),$$

whence

$$f'(z) = nc_n(z-a)^{n-1} + (n+1)c_{n+1}(z-a)^n + \cdots .$$

The logarithmic derivative is

$$\frac{f'(z)}{f(z)} = \frac{1}{z-a} \cdot \frac{nc_n + (n+1)c_{n+1}(z-a) + \cdots}{c_n + c_{n+1}(z-a) + \cdots}$$

where the second factor is a holomorphic function, since $c_n \neq 0$. Thus, by expanding it in its Taylor series,

$$\begin{aligned} \frac{f'(z)}{f(z)} &= \frac{1}{z-a} \cdot \left(n + d_0(z-a) + d_1(z-a)^2 + \cdots\right) \\ &= \frac{n}{z-a} + d_0 + d_1(z-a) + \cdots . \end{aligned}$$

So this yields the Laurent series of the logarithmic derivative $(\log \circ f)'$ in a neighborhood of a. And we see that a is a pole of order 1 with residue n. Therefore:
The roots of a function f are poles of order 1 of its logarithmic derivative $(\log \circ f)'$, and the logarithmic residue of each root equals its multiplicity.

Let f be holomorphic on a region $D \subseteq \mathbb{C}$ (i.e. D is open and connected in $\mathbb{C}$) and continuous on $\bar{D}$. Furthermore, suppose f does not vanish on the boundary C of D, and f' is continuous on C.
Then, there are only finitely many roots of f in D: otherwise the roots would accumulate in a cluster-point a in $\bar{D}$. If a lies in the interior of D, then $f(z)$ vanishes identically on D (identity theorem) and so, by continuity, on $\bar{D}$. If $a \in C$, then $f(a) = 0$ by continuity, in both cases a contradiction.
Assume the roots of f in D have multiplicities $n_1, \ldots, n_l$. Applying the theorem of residues gives

$$\frac{1}{2\pi i} \int_C \frac{f'(z)}{f(z)} dz = n_1 + \cdots + n_l = N,$$

where N is the number (with multiplicities) of roots of f in D.
Observe that, since $f(z) = |f(z)| e^{i \arg f(z)}$,

$$\begin{aligned} \frac{1}{2\pi i} \int_C \frac{f'(z)}{f(z)} dz &= \frac{1}{2\pi i} \int_C d \log f(z) \\ &= \frac{1}{2\pi i} \int_C d \log |f(z)| + \frac{1}{2\pi} \int_C d \arg f(z) \end{aligned}$$

where by log and arg is meant one branch of these functions which is continuously defined along C, respectively. The first integral on the right-hand side vanishes, since $\log |f(z)|$ returns to the starting value by running through whole C. The quantity

$$\frac{1}{2\pi} \int_C d \arg f(z) = \frac{1}{2\pi} \Delta_C \arg f(z)$$

is the increment of $\arg f(z)$, divided by 2π, if z runs through C once in mathematical positive direction. It vanishes, if the origin is not contained in the interior of $f(C)$.

Summarizing:
Suppose f is holomorphic on a region D, continuous on $\bar{D}$, f does not vanish on $C = \partial D$, and f' is continuous on C. Then, for the number N of roots (with multiplicities) of f in D we have

$$N = \frac{1}{2\pi i}\int_C \frac{f'(z)}{f(z)}dz = \frac{1}{2\pi}\Delta_C \arg f(z).$$

Now consider the following corollary:

THEOREM 1.2.1 (Rouché). *Suppose the functions $f, g : \mathbb{C} \to \mathbb{C}$ are holomorphic in the interior of a simple closed curve C, continuous on C, and they satisfy the condition $|f(z)| > |g(z)|$ for $z \in C$. Then, the function $f + g$ has as many roots in the interior of C as f.*

PROOF. We have $|f(z)| > |g(z)| \geq 0$ and $|f(z) + g(z)| \geq |f(z)| - |g(z)| > 0$ if $z \in C$, hence both functions, f and $f + g$, cannot vanish on C. Consider

$$\begin{aligned} \arg\left(\frac{f(z)+g(z)}{f(z)}\right) &= \arg\left(f(z)+g(z)\right) + \arg\left(\frac{1}{f(z)}\right) \\ &= \arg\left(f(z)+g(z)\right) - \arg f(z). \end{aligned}$$

This implies

$$\Delta_C \arg\left(f(z)+g(z)\right) = \Delta_C \arg f(z) + \Delta_C \arg\left(1 + \frac{g(z)}{f(z)}\right).$$

Let now z run through C. Since $\left|\frac{g(z)}{f(z)}\right| < 1$ on C, the point $w(z) = 1 + \frac{g(z)}{f(z)}$ lies in the interior of the circle $\{w : |w-1| = 1\}$. Therefore $w(z)$ cannot run around the origin, whence $\Delta_C \arg\left(1 + \frac{g(z)}{f(z)}\right) = 0$. Consequently,

$$\Delta_C \arg\left(f(z)+g(z)\right) = \Delta_C \arg f(z)$$

from which the statement follows. □

This enables us to prove the following theorem on the continuity of the roots of an equation depending on parameters:

THEOREM 1.2.2. *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 $z_j = z_j(t)$ $(j = 1, \ldots, r)$, and $\lim_{t\to t_0} z_j(t) = z_0$ $(j = 1, \ldots, r)$.*

PROOF. By assumption $z \mapsto f(t_0, z)$ is holomorphic on A, it does not vanish identically, and $z_0 \in A$. Therefore, z_0 is an isolated root of $f(t_0, z)$, and we may choose a small circle C in A with center z_0, such that z_0 is the only root of $f(t_0, z)$ lying in the interior of C, and no root is lying on C.
Let $m = \min_{z\in C} |f(t_0, z)|$, then $m > 0$, since C is compact and $z \mapsto f(t_0, z)$ is continuous. By continuity of f in both variables, for each $z \in C$ there is a neighborhood U_z of z contained in A and a neighborhood V_z of t_0 in $\mathbb{R}$ such that $|f(t, w) - f(t_0, z)| \leq |f(t, w) - f(t, z)| + |f(t, z) - f(t_0, z)| < \frac{m}{2}$ for all $w \in U_z$ and $t \in V_z$. The compact C 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 $z \in C$ and all $t \in V$

$$|f(t, z) - f(t_0, z)| < \min_{z\in C} |f(t_0, z)| \leq |f(t_0, z)|.$$

We can apply Rouché's theorem. Consequently, for $t \in V$ the equation $f(t, z) = 0$ has as many roots in the interior of C 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 C to the point z_0. $\square$

Remark. The parameter space $\mathbb{R}$ in theorem 1.2.2 can be replaced by any metric space.

We shall discuss another continuity result in section 2.4. It will yield a global continuous parameterization of the roots of a hyperbolic polynomial.

CHAPTER 2

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 [**1**]. The last section 2.7 gives a short glance to the complex case, where there are no restrictions on the roots to be real.

2.1. Choosing differentiable square roots

For 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 parameter t. By replacing the variable x with $y = x - \frac{a_1(t)}{2}$, we reduce the problem to $a_1 \equiv 0$.

Proposition 2.1.1. *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 2.3.4), then smooth roots x 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 nonzero 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 $h \geq 0$ 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 2.1.2 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 $f(t) = (t-t_0)^2 h(t)$, 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 $f(t) = 0 < f''(t)$. We put $x(t) = \epsilon(t)y(t)$, where $\epsilon(t) \in \{\pm 1\}$ will be chosen later. The set $\{t : f(t) = f''(t) = f^{(4)}(t) = 0\}$ is closed and has a countable disjoint union of open intervals as complement. In each of these intervals choose a point t_0 with $f(t_0) > 0$, near which y is C^4. Put $\epsilon(t_0) = 1$, and let ϵ change sign exactly at points with $f(t) = f''(t) = 0$ but $f^{(4)}(t) > 0$. These points do not accumulate inside each interval. Therefore, x is twice differentiable, and the proof is complete. □

LEMMA 2.1.2. *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*

$$f'(t)^2 \leq 2f(t) \max\{f''(t_0 + r(t - t_0)) : 0 \leq r \leq 2\}. \tag{2.1}$$

Proof. Since f is non-negative, $f(t) = 0$ implies $f'(t) = 0$, so (2.1) holds at zeros of f. Hence we assume $f(t) > 0$. We use Taylor's formula

$$f(t+s) = f(t) + f'(t)s + \int_0^1 (1-r)f''(t+rs)dr s^2. \tag{2.2}$$

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

$$\begin{aligned} f(t) &= 0 + 0 + \int_0^1 (1-r)f''(t_0 + r(t-t_0))dr(t-t_0)^2 \\ &\leq \frac{(t-t_0)^2}{2} \max\{f''(t_0 + r(t-t_0)) : 0 \leq r \leq 2\}. \end{aligned} \tag{2.3}$$

Now in (2.2) we replace s by $-\epsilon s$, where $\epsilon = \operatorname{sgn}(f'(t))$, to obtain

$$0 \leq f(t - \epsilon s) = f(t) - |f'(t)|s + \int_0^1 (1-r)f''(t - \epsilon rs)dr s^2. \tag{2.4}$$

Let us assume $t \geq t_0$ and put

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

then, by (2.3) and since $f(t) > 0$ (which implies $\max\{f''(t_0 + r(t-t_0)) : 0 \leq r \leq 2\} > 0$), we have $0 < s(t) \leq t - t_0$, and $s(t)$ is well defined. This choice of s in (2.4) gives

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

which proves (2.1) for $t \geq t_0$. Since the assertion is symmetric, it then holds for all t. □

Note that the differentiability assumptions imposed on f in proposition 2.1.1 are best possible: if they are slightly weakened, then the statements are false.
If $f \geq 0$ 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 $f \geq 0$ 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 $f \geq 0$ 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 $f \geq 0$ 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 12.2 (chapter III) in [**20**]: Let $h : \mathbb{R} \to [0,1]$ be smooth with $h(t) = 1$ for $t \geq 0$ and $h(t) = 0$ for $t \leq -1$. Then, we claim that the function

$$f(t) := \sum_{n=1}^{\infty} h_n(t - t_n) \cdot \left(\frac{2n}{2^n}(t - t_n)^2 + \frac{1}{4^n} \right), \tag{2.5}$$

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 (2.5) 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 $h_n(t-t_n) \neq 0$ 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,

$$\begin{aligned}
n^2\sup\{|(h_n\cdot c_n)^{(k)}(t)| : t\in\mathbb{R}\} &= n^2\sup\left\{|(h_n\cdot c_n)^{(k)}(t)| : |t| \le \frac{1}{n\cdot 2^{n+1}}+\frac{1}{n^2}\right\}\\
&\le n^2\sum_{i=0}^{k}\binom{k}{i}n^{2i}H_i\sup\left\{|c_n^{(k-i)}(t)| : |t|\le \frac{1}{n\cdot 2^{n+1}}+\frac{1}{n^2}\right\}\\
&\le \left(\sum_{i=0}^{k}\binom{k}{i}n^{2i+2}H_i\right)\sup\{|c_n^{(j)}(t)| : |t|\le 2, j\le k\}, \qquad (2.6)
\end{aligned}$$

since

$$\begin{aligned}
h_n^{(i)}(t) &= \sum_{j=0}^{i}\binom{i}{j}(-1)^{i-j}n^{2i}\cdot h^{(j)}\left(n^2\left(\frac{1}{n\cdot 2^{n+1}}+t\right)\right)\\
&\qquad \underbrace{\cdot\, h^{(i-j)}\left(n^2\left(\frac{1}{n\cdot 2^{n+1}}-t\right)\right)}_{=1 \text{ for } i=j \text{ and } =0 \text{ for } j<i}\\
&= n^{2i}\cdot h^{(i)}\left(n^2\left(\frac{1}{n\cdot 2^{n+1}}+t\right)\right).
\end{aligned}$$

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 (2.6) is bounded with respect to $n\in\mathbb{N}$. Consequently, the series $\sum_n h_n(\quad -t_n)c_n(\quad -t_n)$ 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 $\lim t_n$.

Note that there are results concerning higher-dimensional parameter spaces: In [**13**] 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, a smooth function $f \geq 0$ is constructed which is flat at 0 such that the positive square root is not C^2. In [**12**] Dieudonné gave shorter proofs of Glaeser's results.

2.2. The space of hyperbolic polynomials

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

There is an elegant description of the space of hyperbolic polynomials with degree n as semialgebraic set in $\mathbb{R}^n$.

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:

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

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

$$s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 + \cdots + (-1)^{k-1} s_1 \sigma_{k-1} + (-1)^k k \sigma_k = 0 \quad (k \geq 1). \tag{2.7}$$

This relation 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 (the determinant of the derivative) of s^n is $n!$-times the Vandermonde determinant:

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

Even the derivative itself $ds^n(x)$ 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

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

Denote by B_k the minor formed by the first k rows and columns of B. From

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

it follows that

$$\Delta_k(x) := \det(B_k(x)) = \sum_{i_1<i_2<\cdots<i_k} (x_{i_1}-x_{i_2})^2\cdots(x_{i_1}-x_{i_k})^2\cdots(x_{i_{k-1}}-x_{i_k})^2, \tag{2.8}$$

since for $n\times k$ - matrices A one has $\det(AA^\top) = \sum_{i_1<\cdots<i_k}\det(A_{i_1,\dots,i_k})^2$, where $A_{i_1,\dots,i_k}$ is the minor of A with indicated rows and columns. 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 an unique symmetric $n\times n$ - matrix $\tilde{B}$ with $B = \tilde{B}\circ\sigma^n$.

The following theorem is Sylvester's version of a theorem of Sturm giving a nice characterization of the space of hyperbolic polynomials of degree n. The proof presented here is due to Procesi [**32**].

THEOREM 2.2.1. *Let P be a monic polynomial of degree n with real coefficients $a_1,\dots,a_n$. Then 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 $\tilde{\Delta}_k(P) = \tilde{\Delta}_k(a_1,\dots,a_n)\geq 0$ for $1\leq k\leq n$.*

Moreover, the rank of B 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 $I := (P(x))$ 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 $\mathrm{tr}: T\to\mathbb{R}$, 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 $\mathrm{tr}(a) := \mathrm{tr}(\bar{a})$. Then, $(a,b) := \mathrm{tr}(ab)$ is a symmetric bilinear form, and we can associate a quadratic form $F(a) := \mathrm{tr}(a^2)$.
Let J be the Jacobson radical of T, i.e., the intersection of all maximal ideals of T, and set $\bar{T} = T/J$. 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 $\beta_{k+1},\bar{\beta}_{k+1},\dots,\beta_{k+s},\bar{\beta}_{k+s}$ the distinct complex roots of P. The trace map tr factors through $\bar{T}$ and gives

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

where m_i is the multiplicity of the root β_i ($1\leq i\leq k+s$).
We assert that the quadratic form $F(a) = \mathrm{tr}(a^2)$ (considered as form on $\bar{T}$) is positive definite if and only if $s = 0$. This can be easily seen from the following formula

$$\begin{aligned} F(a) &= \mathrm{tr}(a^2) = \mathrm{tr}(a_1^2,\dots,a_{k+s}^2) \\ &= \sum_{i=1}^k m_i a_i^2 + \sum_{j=1}^s m_{k+j}(a_{k+j}^2+\bar{a}_{k+j}^2). \end{aligned}$$

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

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

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 $1, \bar{x}, \dots, \bar{x}^{n-1}$ 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. □

2.3. 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.

Let us consider a smooth curve of hyperbolic polynomials

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

Definition 2.3.1. We will say that the smooth curve of polynomials P is *smoothly solvable near* t_0, if there exist n smooth functions $x_1(t), \dots, x_n(t)$ of the parameter t defined near t_0 which are the roots of $P(t)$ for each t.

Note that the problem of smooth solvability of P can be reduced to $a_1 \equiv 0$, 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.

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 2.3.2. *Let P be a smooth curve of hyperbolic polynomials as above whose roots are all distinct at $t = 0$. 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. □

When there are multiple roots of $P(0)$, we have to invest more effort. A first step playing a key role in the further considerations is the following lemma.

Lemma 2.3.3 (Splitting Lemma). *Let $P_0 = x^n - a_1x^{n-1} + \dots + (-1)^n a_n$ 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

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

where $p + q = n$. Let $x_1, \ldots, x_n$ 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 $(a_1, \ldots, a_n) = \Phi^{p,q}(b_1, \ldots, b_p, c_1, \ldots, c_q)$ 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 $b = (b_1, \ldots, b_p)$ and $c = (c_1, \ldots, c_q)$. From section 2.2 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). □

Now we want to introduce the notion of multiplicity of a function which we shall need when we factorize a curve of polynomials.

DEFINITION 2.3.4. 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 $t \mapsto f(P(t))$.

The splitting lemma 2.3.3 shows that for the problem of smooth solvability it is enough to assume that all roots of $P(0)$ are equal.

PROPOSITION 2.3.5. *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 $t \mapsto x_i(t)$ $(1 \le i \le n)$, and that all roots of $P(0)$ are equal. Then, for all $2 \le k \le n$ we have

$$m(\tilde{\Lambda}_k) \ge k(k-1) \min_{1 \le i \le n} m(x_i)$$

and

$$m(a_k) \ge k \min_{1 \le i \le n} m(x_i).$$

This result holds in the real analytic case and in the holomorphic case, too.

PROOF. The second inequality stated in the proposition follows from $a_k(t) = \sigma_k(x_1(t), \ldots, x_n(t)) = \sum_{1 \le j_1 < \cdots j_k \le n} x_{j_1}(t) \cdots x_{j_k}(t)$. Observe that in equation (2.8) 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. □

LEMMA 2.3.6. *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 (2.7) 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. Since they 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.

LEMMA 2.3.7 (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 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* $2 \leq k \leq n$*;*
(3) $m(a_2) \geq 2r$.

PROOF. We only have to treat $r > 0$.
(1) ⇒ (2): From (2.7) 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): It 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 $a_2(0) = 0$ (because $r > 0$) and lemma 2.3.6 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. This means that $m(a_k) \geq 1$ for $3 \leq k \leq n$. 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 $3 \leq k \leq n$, 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 $m_k \geq kr$.
Let us suppose indirectly that for some $k > 2$ we have $m_k = m(a_k) < kr$. We put

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

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

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

It is easy to see that $\bar{P}_m(t)(x) = t^{-nm}P(t)(t^m x)$, for $t > 0$. So, if $x_1, \dots, x_n$ are the real roots of $P(t)$, then $t^{-m}x_1, \dots, t^{-m}x_n$ 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 2.2.1 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 2.3.6, 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.

Lemma 2.3.8 (Strong Multiplicity Lemma). *Consider a curve of hyperbolic polynomials*

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

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 2.3.6 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 $m_k < k$. In the same way as before (with $r = 1$) we define $m < 1$ and the continuous curve of polynomials $\bar{P}\_m$. 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 it has to hold $m_k = m(a_k)$ for these k, a contradiction. $\square$

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 2.3.9. Consider a smooth curve of hyperbolic polynomials

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

The algorithm consists of following steps:

(1) If all roots of $P(0)$ are pairwise different, P is smoothly solvable for t near 0, by proposition 2.3.2.
(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 2.3.3. 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 $a_1 \equiv 0$ 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 2.2.1 the hyperbolicity of P forces a_2 to be non-positive everywhere: $0 \le \tilde{\Delta}_2 = -2na_2$. We put $m(a_2) := 2r$ for a positive integer r, and from the multiplicity lemma 2.3.7 we obtain $a_k(t) = t^{kr} a_{k,kr}(t)$ near 0 for smooth $a_{k,kr}$ and $2 \le k \le n$. 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 $P_r(t)(x) = t^{-nr} P(t)(t^r x)$, P_r is again a curve of 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 $a_2 \equiv 0$, then all roots of P are identically 0 by lemma 2.3.6, and thus P is smoothly solvable.

(3c) Finally, if $m(a_2)$ is infinite and $a_2 \not\equiv 0$, then by the multiplicity lemma 2.3.7 all $m(a_k)$ for $2 \le k \le n$ 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 $a_1 \equiv 0$. This means that all roots of $P(t)$ meet of infinite order of flatness (see definition 2.3.4) 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 2.4.1, 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 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. 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 f has the form $f(t) = g(t)^2$, then $P^{(\infty)}(t)(x) = x^2 - f(t)$ is smoothly solvable near 0. For the smooth function f defined by (2.5) it is not smoothly solvable.

2.4. Continuous parameterization of the roots

In this section we shall present another continuity result concerning roots of polynomials. It reaches further as the results in chapter 1 in the sense that the following proposition yields even a global continuous parameterization of the roots of hyperbolic polynomials.

Proposition 2.4.1. *For a hyperbolic polynomial*

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

let $x_1(P) \le x_2(P) \le \cdots \le x_n(P)$ be the roots of P, increasingly ordered.
Then all roots $x_i : \sigma^n(\mathbb{R}^n) \to \mathbb{R}$ are continuous.

Proof. First we show that x_1 is continuous. Consider an arbitrary $P_0 \in \sigma^n(\mathbb{R}^n)$. We have to show that for every $\epsilon > 0$ there exists some $\delta > 0$ such that for all $|P - P_0| < \delta$ 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 $x(P) > x_1(P) - \epsilon$ (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 2.3.3 for the C^0-case 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 $P_1(x) = 0$ we obtain

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

whence

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

provided that $n_1 |a_k(P_1)| < \epsilon^k$, which is true for P_1 near P_0, since $a_k(P_0) = 0$ 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 following relations between the coefficients $a_1, \ldots, a_n$ of P and those of $P_2(P)$, 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 $P_2(P)$ are again continuous and so we can proceed by induction on the degree of P. Hence the proposition is proved. □

2.5. Choosing roots of polynomials differentiably

Here we use the results obtained in section 2.3 to construct global smooth roots, if a certain genericity condition ((1) or equivalently (2) in theorem 2.5.1) is fulfilled, and global differentiable roots always. The obstructions contained in the mentioned genericity condition arise in a natural way from the algorithm 2.3.9.

THEOREM 2.5.1. *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 $\tilde{\Delta}_k(P)$ 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 2.3.9 never leads to step (3c).

(3) ⇒ (1): Suppose for contradiction that two roots with different germs at $t = 0$ out of the increasingly ordered continuous roots meet of infinite order at $t = 0$. Then in each application of step (2) in algorithm 2.3.9 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 (2.8) 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 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, by theorem 2.2.1. We assume indirectly that the algorithm 2.3.9 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 $k^{(\infty)} > 2$ 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 in which exactly the k different roots off 0, mentioned at the beginning of this paragraph, appear. Hence, this summand involves exactly $k^{(\infty)}$ many 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. The first part of the proof does not only show that condition (1) and condition (2) in the theorem are equivalent, but also that the algorithm 2.3.9 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, since step (3c) never occurs. Thus we may glue the local solutions to a global solution. This completes the proof. □

THEOREM 2.5.2. *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 coefficients parameterize the roots of P.

PROOF. We follow one step of the algorithm 2.3.9. Without loss of generality we may assume that $a_1 \equiv 0$: replace x by $y = x - \frac{a_1(t)}{n}$ and note that a_1 is C^1. 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. For if it vanished only of first order, then $\tilde{\Delta}_2(P(t)) = -2na_2(t)$ would change sign at $t = 0$, contrary to the assumption that $P(t)$ is hyperbolic for all t, by theorem 2.2.1. Thus $a_2(t) = t^2 a_{2,2}(t)$ near 0 for a continuous function $a_{2,2}$, since $a_2 \in C^2$. By the strong multiplicity lemma 2.3.8, 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 $z_1(t) \le \dots \le z_n(t)$ be its continuous roots, by theorem 2.4.1. Then, $x_j(t) := t \cdot z_j(t)$, where $1 \le j \le n$, are all roots of P, 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 $y_1(t) \le \dots \le y_n(t)$ 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 afterward 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 2.3.3, 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 $y_1(t) \le \dots \le y_n(t)$. 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. □

2.6. The real analytic case

The algorithm 2.3.9 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 2.3.9 cannot occur.
So let $P(t)(x) = x^n - a_1(t)x^{n-1} + \dots + (-1)^n a_n(t)$ be a curve of hyperbolic polynomials, where all $a_i(t)$ are real analytic in t. In analogy to definition 2.3.1, where smooth solvability was defined, we shall 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 are roots of $P(t)$ for all t.

THEOREM 2.6.1. *Let P be a real analytic curve of hyperbolic polynomials*

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

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

Proof. We first show that P is locally real analytically solvable 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 $a_1 \equiv 0$.
The proof will be carried out by induction on the polynomial degree n. If $n = 1$, then the theorem 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 and zero, so by the splitting lemma 2.3.3 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 $a_2 \equiv 0$, then by lemma 2.3.6 all roots of P are identically equal to 0 and we are done. Otherwise, for the multiplicity of the real analytic function a_2 at 0 we have $1 \leq m(a_2) < \infty$, and, again by lemma 2.3.6, 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 2.2.1. So we write $m(a_2) = 2r$ for a positive integer r. Then by the multiplicity lemma 2.3.7 we have $a_k(t) = t^{kr} a_{k,kr}(t)$ for real analytic $a_{k,kr}$ for all $2 \leq k \leq n$. 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 analytic solvable and $x_j(t)$ $(j = 1, \ldots, n)$ are its real analytic roots, then $t^r x_j(t)$ $(j = 1, \ldots, n)$ are the roots of $P(t)$ and so the original curve P is real analytical solvable, too. Now $a_{2,2r}(0) \neq 0$ and we are done by the case above. This shows the claim on local 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. □

Note that the local existence part of this theorem is due to Rellich [**35**], Hilfssatz 2. His proof uses Puiseux-expansions.

Remark. The uniqueness statement of theorem 2.6.1 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.
Theorem 2.6.1 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 above arguments. For $a_1 \neq a_2$ the coefficients describe a real analytic embedding for (t_1,t_2) near 0.

2.7. 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.

Note that the preliminaries presented in section 2.2, including the definition of the Bezoutiant B, its principal minors Δ_k and formula 2.8, are still valid in the present case, where coefficients and roots are complex valued. But there are no restrictions on the coefficients, whence the space of polynomials of degree n to be studied here may be identified with $\mathbb{C}^n$.

As at the beginning of this chapter 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 (2.5) there still not exists a smooth square root.

Note that proposition 2.3.2 and the splitting lemma 2.3.3 are true in the complex case. Also proposition 2.3.5 remains valid, since it follows from (2.8). Evidently, lemma 2.3.6 does not hold anymore and the multiplicity lemma 2.3.7 keeps valid only partially:

Lemma 2.7.1 (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 $r > 1$.
(1) $\Rightarrow$ (2): Exactly the same arguments as in the proof of the multiplicity lemma 2.3.7 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 (2.7) we have $a_k(0)=0$ for all $2 \leq k \leq n$. The rest of the proof coincides with the the one of the multiplicity lemma 2.3.7. □

The proof shows that the multiplicity lemma 2.3.7 holds only partially by the lack of lemma 2.3.6.

As in section 2.3 we may construct an algorithm which extracts the solvable part from the original curve P:

ALGORITHM 2.7.2. 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 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 2.3.2.

(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 2.3.3. 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 $a_1 \equiv 0$ 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, $a_k(0) = 0$ 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 2.3.5. 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

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

If $P_r(t)$ is smoothly (real analytically, holomorphically) solvable and $x_j(t)$ are its smooth (real analytic, holomorphic) roots, then $t^r x_j(t)$ 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 $P_r(0)$ 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.

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 $P(t)(x) = 0$ 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 [**4**].

CHAPTER 3

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 [**8**].

3.1. Introduction: degree 3

To get an idea, how Bronshtein proves the local boundedness of the derivatives of roots of hyperbolic polynomials, we want to discuss the case when the polynomials have degree 3. It will shorten and simplify the general proof essentially but use 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 2.1.1.

Step 1. Let us consider a curve P of monic polynomials of degree 3, i.e., $P(t)(x) = x^3 + A_1(t)x^2 + A_2(t)x + A_3(t)$, having only real roots for all $t \in [-1, 1]$. Assume that the coefficients satisfy $A_i \in C^i([-1, 1])$, for $i = 1, 2, 3$, $A_2(0) \neq 0$ and $A_3(0) = 0$. We want to show that there exists a positive constant C such that

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

To shorten notation let us introduce $a_i = A_i(0)$, for $i = 1, 2, 3$. We define following numbers

$$M_0 = 27\left(\max_{i,t,j\leq i} |A_i^{(j)}(t)| + 4\right)^6 \qquad \text{and} \qquad M_i = M_{i-1}^{100} \quad (i = 1, 2, 3).$$

Their exact value is not really important, the thing that counts is that they are chosen large enough for the estimates to come.
$P(t)(x)$ having all roots real, implies that the same holds for $\frac{\partial}{\partial x}P(t)(x) = 3x^2 + 2A_1(t)x + A_2(t)$, see lemma 3.4.4. And this is equivalent to

$$A_2(t) \leq \frac{1}{3}A_1^2(t). \tag{3.1}$$

Consider the following two cases separately:

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

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 $|t| \leq M_1^{-1}|a_1|$:

$$\frac{1}{2} \leq 1 - \frac{|A_1^{(1)}(\xi)|}{|a_1|}|t| \leq \frac{A_1(t)}{a_1} \leq 1 + \frac{|A_1^{(1)}(\xi)|}{|a_1|}|t| \leq 2, \tag{3.2}$$

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

$$\begin{aligned}
|A_2(t)| &= \left| a_2 + a_2^{(1)} t + \frac{A_2^{(2)}(\xi)}{2!} t^2 \right| \leq |a_2| + |a_2^{(1)}||t| + \frac{|A_2^{(2)}(\xi)|}{2!}|t|^2 \\
&\leq a_1^2 + M_0^2 |a_1| M_1^{-1} |a_1| + M_0 M_1^{-2} a_1^2 \leq 2a_1^2.
\end{aligned} \tag{3.3}$$

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 $|t| \leq \frac{1}{2} M_1^{-1} \left|\frac{a2}{a1}\right| \left(\overset{(3.3)}{\leq} M_1^{-1}|a_1|\right)$,

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

implying

$$\frac{1}{2} \leq \frac{A_2(t)}{a_2} \leq 2, \tag{3.4}$$

for $|t| \leq M_2^{-1} \left|\frac{a2}{a1}\right|$.

For a root $x(t)$ of $P(t)$, 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 $\frac{\partial}{\partial x} P(t)(x)$ 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 of $P(t)$ lying between $x_1(t)$ and $x_2(t)$ (see lemma 3.4.4). Therefore,

$$\begin{aligned}|A_3(t)| \quad &\leq \quad 64\left(|A_1(t)| + \left|\frac{A_2(t)}{A_1(t)}\right|\right)^3 + 16|A_1(t)|\left(|A_1(t)| + \left|\frac{A_2(t)}{A_1(t)}\right|\right)^2 \\ &\quad + 4|A_2(t)|\left(|A_1(t)| + \left|\frac{A_2(t)}{A_1(t)}\right|\right). \end{aligned} \tag{3.5}$$

We know, by (3.2) and (3.3), that, for $|t| \leq M_1^{-1}|a_1|$, $|A_1(t)| \leq 2|a_1|$ and $\left|\frac{A_2(t)}{A_1(t)}\right| \leq 4|a_1|$. Hence, using this to estimate the right-hand side of (3.5),

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

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

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

Use lemma 3.4.2 to get following estimates

$$|a_3^{(j)}| \leq M_0M_1^{2+j}|a_1|^{3-j} \qquad (j = 0, 1, 2). \tag{3.6}$$

Therefore,

$$|A_3''(t)| = |a_3^{(2)} + A_3^{(3)}(\xi)t| \leq M_0M_1^4|a_1| + M_0M_1^{-1}|a_1| \leq M_0^2M_1^4|a_1|, \tag{3.7}$$

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

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 for one t. We claim that one of the roots has the form $-q\frac{A_2(t)}{A_1(t)}$ with $0 < q \leq 1$.

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 $> -\frac{A_2(t)}{A_1(t)} = \frac{2x_1(t)x_2(t)}{x_1(t)+x_2(t)}$ (otherwise $-\frac{A_2(t)}{A_1(t)} < 0$). If this holds for both, $x_1(t)$ and $x_2(t)$, then $x_1(t) > x_2(t)$ and $x_1(t) < x_2(t)$ follow, 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 $-q\frac{A_2(t)}{A_1(t)}$ $(0 < q \leq 1)$ with minimal absolute-value. If $x_0(t) \neq 0$ then $\frac{\partial}{\partial x}P(t)(x)$ has constant sign on the open segment between 0 and $x_0(t)$. Therefore, $0 \geq \frac{\partial^2}{\partial x^2}P(t)(x_0(t)) \cdot x_0(t) \cdot \frac{\partial}{\partial x}P(t)(0) = \frac{\partial^2}{\partial x^2}P(t)(x_0(t)) \cdot \left(-q\frac{A_2(t)}{A_1(t)}\right) \cdot A_2(t)$, implying $0 \leq \frac{\partial^2}{\partial x^2}P(t)(x_0(t)) \cdot A_1(t)$. If $x_0(t) = 0$ we come to the same result, since then $\frac{\partial^2}{\partial x^2}P(t)(x_0(t)) \cdot A_1(t) = 2A_1^2(t) \geq 0$.

We want to use these consultations to find an estimate for $A_3(t)A_1(t)$. For $A_1(t) = 0$, it is trivial. Thus, suppose $A_1(t) \neq 0$ and consider the following cases: If $x_0(t)$ is a root of $P(t)$, then

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

If $x_0(t)$ is not a root of $P(t)$ then $P(t)(x_0(t)) \neq 0$ and $\frac{\partial^2}{\partial x^2}P(t)(x_0(t)) \neq 0$ (giving the curvature) have to have different signs (for details see lemma 3.4.4). Assume that $\frac{\partial^2}{\partial x^2}P(t)(x_0(t)) > 0$. Then 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 a 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. But 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 (3.2), (3.3) and (3.4), we can conclude that, for $|t| \leq M_1^{-1}|a_1|$, this estimate gives:

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

In this inequality we plug $t = \pm M_1^{-3}\left|\frac{a_2}{a_1}\right|$ (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 calculate as follows:

$$\begin{aligned} \pm a_1 a_3^{(1)} M_1^{-3}\left|\frac{a_2}{a_1}\right| &\leq 320a_2^2 - a_1\frac{A_3^{(2)}(\xi)}{2!}M_1^{-6}\left|\frac{a_2}{a_1}\right|^2 \\ &\leq 320a_2^2 + |a_1|\frac{|A_3^{(2)}(\xi)|}{2!}M_1^{-6}\left|\frac{a_2}{a_1}\right|^2 \\ &\overset{(3.7)}{\leq} 320a_2^2 + \frac{1}{2}M_0^2 M_1^{-2}a_2^2 \\ &\leq M_0 a_2^2. \end{aligned}$$

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 (3.1):

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

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

$$\frac{1}{2} \le 1 - \frac{|a_2^{(1)}|}{|a_2|}|t| - \frac{|A_2^{(2)}(\xi)|}{2|a_2|}|t|^2 \le \frac{A_2(t)}{a_2} \le 1 + \frac{|a_2^{(1)}|}{|a_2|}|t| + \frac{|A_2^{(2)}(\xi)|}{2|a_2|}|t|^2 \le 2, \quad (3.8)$$

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

$$|A_1(t)| = |a_1 + A_1^{(1)}(\xi)t| \le |a_1| + |A_1^{(1)}(\xi)||t| \le |a_2|^{\frac{1}{2}} + M_0 M_2^{-1}|a_2|^{\frac{1}{2}} \le 2|a_2|^{\frac{1}{2}},$$

and

$$\left|\frac{A_2(t)}{A_1(t)}\right| \overset{(3.1)}{\le} \frac{1}{3}|A_1(t)|.$$

Apply these estimates to (3.5):

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

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

$$\begin{aligned} \left|a_3 + a_3^{(1)}t + \frac{a_3^{(2)}}{2!}t^2\right| &= \left|A_3(t) - \frac{A_3^{(3)}(\xi)}{3!}t^3\right| \le |A_3(t)| + \frac{|A_3^{(3)}(\xi)|}{3!}|t|^3 \\ &\le M_0|a_2|^{\frac{3}{2}} + M_0 M_2^{-3}|a_2|^{\frac{3}{2}} \le M_0^2|a_2|^{\frac{3}{2}}, \end{aligned}$$

for $|t| \le M_2^{-1}|a_2|^{\frac{1}{2}}$. By using lemma 3.4.2 we get

$$|a_3^{(j)}| \le M_0 M_2^{2+j}|a_2|^{\frac{3-j}{2}} \qquad (j = 0, 1, 2) \tag{3.9}$$

which, take $j = 1$, concludes case (B).

Step 2. Note that, if in the assumtions of step 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 (3.6). Case (B) does not appear, since $a_2 = 0$ would imply $a_1 = 0$, contrary to the assumption.

Step 3. We will need similar results to those in step 1 and step 2 also for the degrees one and two. But these are more easily to get: For $P(t)(x) = x + A_1(t)$, where $A_1 \in C([-1, 1])$, 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 $A_2(0) = 0$. Set $t = \pm M_0^{-1}a_1$ (let $M_0 = 4(\max_{i,t,j\le i}|A_i^{(j)}(t)| + 4)^4$ here) 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,

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

whence

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

Estimates of the kind as in step 2 are trival for degree one and two.

Step 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 $i = 1,2,3$. We assert that for any compact $K \subset (-1,1)$ there exists a constant C_K such that all roots x_j $(j = 1,2,3)$ 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 $K \subset (-1,1)$ for a $j \in \{1,2,3\}$. Without loss of generality, say $j = 1$, and assume there is a sequence $(t_p)_{p\in\mathbb{N}}$ in K such that $t_p \overset{p\to\infty}{\longrightarrow} t_\infty$, $x_1(t_p) \overset{p\to\infty}{\longrightarrow} x_1(t_\infty)$ and $|x_1'(t_p)| \overset{p\to\infty}{\longrightarrow} \infty$. By switching to a subsequence, we can achieve that $x_1(t_p)$ has fix multiplicity q for all $p \in \mathbb{N}$, and $x_1(t_\infty)$ has multiplicity $s \geq q$. 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 $b_{p,0} \equiv 1$. As we will see in theorem 3.2.1, $x_1'(t_p)$ has to satisfy 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 $x_1'(t_p)$ were bounded (see lemma 3.4.3), and we were done.
Observe that

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

and

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

since $x_1(t_p)$ has multiplicity q. 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 $q \geq j$. Note that this polynomial has at most degree three, all of its roots are always real, and all coefficients are of class C^3. Let us apply step 1 and step 2 to it, for $j=1$ and $j=2$, respectively. We find that

$$\left|\frac{b_{p,3-q+j}^{(j)}(t_p)}{b_{p,3-q}(t_p)}\right| \le C \qquad (j=1,2),$$

where C does not depend on p, since $t_p \in K$. If $q<3$ we are done. If $q=3$, then

$$\left|\frac{b_{p,3}^{(3)}(t_p)}{b_{p,0}(t_p)}\right| = |b_{p,3}^{(3)}(t_p)|$$

is also bounded, since $b_{p,3}^{(3)}$ is continuous and $t_p \in K$. This shows the assertion.

3.2. 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, compare with theorem 2.5.2. Note that this approach provides a polynomial equation the potential derivatives of a root have to satisfy, namely equation (3.10).

THEOREM 3.2.1. *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 $i=1,\dots,n$. Then at any point $t_0 \in (-1,1)$ all the roots $x_j = x_j(t)$ $(j=1,\dots,n)$ of P (with suitable choice of the branches) are differentiable.
Moreover, each of the q possible derivatives at t_0 of a q-fold root of $P(t_0)$ satisfies the following hyperbolic equation:

$$b_0^{(0)}(t_0)x^q + \frac{1}{1!}b_1^{(1)}(t_0)x^{q-1} + \cdots + \frac{1}{q!}b_q^{(q)}(t_0) = 0, \tag{3.10}$$

where

$$b_i(t) = \frac{1}{(q-i)!}\left(\frac{\partial}{\partial x}\right)^{q-i}\Bigg|_{x=x(t_0)} P(t)(x) \qquad (i=0,\dots,q).$$

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

$$\begin{aligned} Q(t)(x) &= P(t)(x+x_0) \\ &= \sum_{j=0}^{n-q-1} \tilde a_j(t)x^{n-j} + \sum_{j=0}^{q} b_j(t)x^{q-j} \\ &= \tilde a_0(t)x^n + \cdots + \tilde a_{n-q-1}(t)x^{q+1} + b_0(t)x^q + \cdots + b_q(t). \end{aligned} \tag{3.11}$$

Note that $\tilde a_0 = 1$. The coefficients $\tilde a_j$ are again of class C^k on $(-1,1)$ and (compare with Taylor's formula)

$$b_j(t) = \frac{1}{(q-j)!}\left(\frac{\partial}{\partial x}\right)^{q-j}\Bigg|_{x=x_0} P(t)(x) \qquad (j=0,\dots,q). \tag{3.12}$$

Of course, $\tilde a_0, \dots, \tilde a_{n-q-1}$ are Taylor-coefficients, too, but we are not interested in their explicit form. Put

$$b_j^{(i)} = \left(\frac{d}{dt}\right)^i\Bigg|_{t=0} b_j(t) \qquad (0 \le i,j \le q).$$

Since x_0 is a q-fold root of $P(0)$, we find $b_0^{(0)} \neq 0$ and $b_j^{(0)} = 0$, for $j=1,\dots,q$.

Claim 1. For $j = 1, \dots, q$ and $i = 0, \dots, j-1$, we have $b_j^{(i)} = 0$.

Note that 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 $\{1, \dots, q\}$ for which claim 1 is not true. Thus,

$$\exists i_0 \in \{0, \dots, j_0 - 1\} : \quad b_{j_0}^{(i)} = 0 \quad \text{for each} \quad i < i_0 \quad \text{and} \quad b_{j_0}^{(i_0)} \neq 0.$$

Consider the polynomial $|t|^{-i_0}(\frac{\partial}{\partial x})^{q-j_0} Q(t)(x)$ and replace x by $|t|^{\frac{i_0}{j_0}} \tilde{x}$. Since differentiating a polynomial in x with respect to x leaves invariant hyperbolicity (see lemma 3.4.4(1)), the resulting polynomial $R(t)(\tilde{x})$ is hyperbolic, and it takes the following form

$$\begin{aligned} R(t)(\tilde{x}) \quad &= \quad \sum_{j=0}^{n-q-1} \frac{(n-j)!}{(n-q-j+j_0)!} \tilde{a}_j(t) |t|^{\frac{i_0}{j_0}(n-q-j)} \tilde{x}^{n-q-j+j_0} \\ &\quad + \sum_{j=0}^{j_0} \frac{(q-j)!}{(j_0-j)!} b_j(t) |t|^{-\frac{i_0}{j_0} j} \tilde{x}^{j_0 - j}. \end{aligned} \tag{3.13}$$

Analyze the coefficients in the second sum of (3.13):
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,

$$\begin{aligned} b_j(t)|t|^{-\frac{i_0}{j_0} j} \quad &= \quad \left(b_j^{(0)} + \dots + \frac{b_j^{(j-1)}}{(j-1)!} t^{j-1} + \frac{b_j^{(j)}(\xi)}{j!} t^j \right) |t|^{-\frac{i_0}{j_0} j} \\ &= \quad \frac{1}{j!} b_j^{(j)}(\xi) \cdot \operatorname{sgn}\left(t^{-\frac{i_0}{j_0} j} \right) \cdot t^{j - \frac{i_0}{j_0} j} \end{aligned}$$

is continuous in t and takes the value 0 for $t = 0$.
For $j = j_0$, we get

$$\begin{aligned} b_{j_0}(t)|t|^{-i_0} \quad &= \quad \left(b_{j_0}^{(0)} + \dots + \frac{b_{j_0}^{(i_0)}}{i_0!} t^{i_0} + \frac{b_{j_0}^{i_0+1}(\xi)}{(i_0+1)!} t^{i_0+1} \right) |t|^{-i_0} \\ &= \quad \frac{b_{j_0}^{(i_0)}}{i_0!} \cdot \operatorname{sgn}\left(t^{-i_0} \right) + \frac{b_{j_0}^{i_0+1}(\xi)}{(i_0+1)!} \cdot \operatorname{sgn}\left(t^{-i_0} \right) \cdot t, \end{aligned}$$

where the second term is again continuous in t and vanishes for $t = 0$.
Clearly, the coefficients in the first sum of (3.13) are continuous in t and vanish for $t = 0$, too. Thus, (3.13) can be written as follows:

$$\begin{aligned} R(t)(\tilde{x}) \quad &= \quad \frac{q!}{j_0!} b_0^{(0)} \tilde{x}^{j_0} + \frac{(q-j_0)!}{i_0!} b_{j_0}^{(i_0)} \operatorname{sgn}\left(t^{i_0} \right) \\ &\quad + c_0(t) \tilde{x}^{n-q+j_0} + \dots + c_{n-q+j_0}(t), \end{aligned} \tag{3.14}$$

where $c_0, \dots, c_{n-q+j_0}$ are continuous functions in t, and all of them vanish for $t = 0$. Before we can 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 $0 < i_0 < j_0 \geq 2$, has non-real roots, for $t b_0^{(0)} b_{j_0}^{(i_0)} > 0$.

If j_0 is odd, then $j_0 \geq 3$, and the equation has non-real roots whenever $t \neq 0$, in particular, when $t b_0^{(0)} b_{j_0}^{(i_0)} > 0$. 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 $j_0 \geq 4$ can be reduced to the considered two cases, by substitution. Thus, claim 2 is proved.

Now, consider the polynomial $R(t)(\tilde{x})$ in (3.14). For t near 0 such that the condition $tb_0^{(0)}b_{j_0}^{(i_0)} > 0$ in claim 2 is satisfied, theorem 1.2.2 tells us that the polynomial $R(t)(\tilde{x})$ has non-real roots, a contradiction. Thus, claim 1 follows.

Putting $x = t\tilde{x}$ in equation (3.11) 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 $j = 0, \ldots, q$,

$$\begin{aligned} b_j(t)t^{-j} &= \left(b_j^{(0)} + \cdots + \frac{b_j^{(j)}}{j!}t^j + O(t^j)\right)t^{-j} \\ &= \frac{b_j^{(j)}}{j!} + t^{-j}O(t^j), \end{aligned}$$

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 $j = 0, \ldots, n$. We use again theorem 1.2.2, and 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

$$0 = t^{-q}Q(t)(t\tilde{x}_j(t)) = t^{-q}P(t)(x_0 + t\tilde{x}_j(t)) \qquad (j = 1, \ldots, q)$$

implies that, for t near 0, $P(t)$ has q roots of the form $x_j(t) = x_0 + t\tilde{x}_j(t)$, with $j = 1, \ldots, q$. 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

$$b_i^{(i)} = \frac{1}{(q-i)!}\left(\frac{\partial}{\partial t}\right)^i\Big|_{t=0}\left(\frac{\partial}{\partial x}\right)^{q-i}\Big|_{x=x(0)} P(t)(x) \qquad (i = 0, \ldots, q).$$

Therefore, the theorem is proved. □

3.3. 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 parameterized differentiably on the one hand (theorem 2.5.2), with those Bronshtein uses on the other hand (theorem 3.2.1). It will turn out that the two approaches are very similar with the decisive difference that the iterative method in the proof of theorem 2.5.2 following the algorithm 2.3.9 is done simultaneously by Bronshtein.

In the following we shall repeat the main steps in Bronshtein's proof of theorem 3.2.1 and comment them from Alekseevsky, Kriegl, Losik and Michor's point of view. The main ingredients of their proof of theorem 2.5.2 are the splitting lemma

2.3.3, the multiplicity lemma 2.3.8, and proposition 2.4.1 providing 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 continuous roots $x_1(t), \dots, x_q(t)$ for t near 0 which agree for $t = 0$ and are differentiable there. Here implicitly is used the splitting lemma 2.3.3. 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

$$\begin{aligned} Q(t)(x) &= P(t)(x + x_0) \\ &= \tilde{a}_0(t)x^n + \cdots + \tilde{a}_{n-q-1}(t)x^{q+1} + b_0(t)x^q + \cdots + b_q(t), \end{aligned}$$

where

$$b_j(t) = \frac{1}{(q-j)!} \left(\frac{\partial}{\partial x} \right)^{q-j} \Big|_{x=x_0} P(t)(x) \qquad (j = 0, \dots, q),$$

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 $x \rightsquigarrow x + \frac{a_1(t)}{n}$ or the assumption $a_1 \equiv 0$ in the proof of theorem 2.5.2.

Claim 1 states that $b_j^{(i)}(0) = 0$, for $j = 1, \dots, q$ and $i = 0, \dots, j-1$, 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 2.3.8.

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 $j = 0, \dots, n$. This continuous curve of hyperbolic polynomials $t^{-q}Q(t)(t\tilde{x})$ corresponds to the curve $P_1(t)(x)$ in the proof of theorem 2.5.2. 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 1.2.2 and proposition 2.4.1, respectively, such that $x_j(t) = x_0 + t\tilde{x}_j(t)$ $(j = 1, \dots, q)$ 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 3.5.3, 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 unique 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 2.3.3 such this P is not the curve of polynomials we have started from). Since we have applied the splitting lemma 2.3.3, $P_1(t)$ 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.

3.4. Estimating coefficients of hyperbolic polynomials

In this section are collected the preliminaries used in section 3.5. They consist mostly of estimates of the coefficients of hyperbolic polynomials.

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 the so-called *Vieta's formulas*

$$a_i = \sum_{1 \le j_1 < \cdots < j_i \le n} x_{j_1} \cdots x_{j_i} \qquad (i = 1, \dots, n).$$

So we see that the coefficients of P are (up to their sign) the elementary symmetric functions in its roots.

LEMMA 3.4.1. *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 of all let us assume that a_{n-1}, a_{n-2} and a_{n-3} do not vanish. With this assumption consider the following two cases:

(1) $2n|x_1| \ge |x_2|$: From Vieta's formulas we have:

$$|a_n| = |x_1 x_2 \cdots x_n|,$$

$$\begin{aligned} |a_{n-1}| &= \left| \sum_{1 \le j_1 < \cdots < j_{n-1} \le n} x_{j_1} \cdots x_{j_{n-1}} \right| \\ &\le \sum_{1 \le j_1 < \cdots < j_{n-1} \le n} |x_{j_1} \cdots x_{j_{n-1}}| \\ &\le n |x_2 \cdots x_n| \end{aligned}$$

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, by our assumption that $a_{n-1} \ne 0$, 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) $2n|x_1| < |x_2|$: With Vieta's formulas we find:

$$|a_{n-1}| = \left|(x_2 \cdots x_n)\left(1 + \frac{x_1}{x_2} + \cdots + \frac{x_1}{x_n}\right)\right|,$$

$$|a_{n-2}| \leq \frac{n(n-1)}{2}|x_3 \cdots x_n|$$

and

$$|a_{n-3}| \leq \frac{n(n-1)(n-2)}{6}|x_4 \cdots x_n|.$$

Consider

$$\begin{aligned}\left|1 + \frac{x_1}{x_2} + \cdots + \frac{x_1}{x_n}\right| &= \left|1 - \left(-\frac{x_1}{x_2} - \cdots - \frac{x_1}{x_n}\right)\right| \\ &\geq \left|1 - \left|\frac{x_1}{x_2} + \cdots + \frac{x_1}{x_n}\right|\right| \\ &> \frac{1}{2},\end{aligned}$$

since

$$\left|\frac{x_1}{x_2} + \cdots + \frac{x_1}{x_n}\right| \leq \left|\frac{x_1}{x_2}\right| + \cdots + \left|\frac{x_1}{x_n}\right| < (n-1)\frac{1}{2n} < \frac{1}{2}.$$

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

$$|x_2|^2 \leq |x_2 x_3| = \frac{|x_2 \cdots x_n|}{|x_4 \cdots x_n|} < \frac{2|a_{n-1}|}{|a_{n-3}|} \cdot \frac{n(n-1)(n-2)}{6} < n^4 \left|\frac{a_{n-1}}{a_{n-3}}\right|.$$

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 ∞, so the inequality is 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 $2n|x_1| \geq |x_2|$, 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 $2n|x_1| \geq |x_2|$, the statement is trivial. If $2n|x_1| < |x_2|$, investigate again 2.: $a_{n-1} = 0$ would imply that 0 is a 3-fold root; for $a_{n-1} \neq 0$, the inequality $|x_2| \leq n^2 \left|\frac{a_{n-1}}{a_{n-2}}\right|$ is clear and $|x_2| \leq n^2 \left|\frac{a_{n-1}}{a_{n-3}}\right|^{\frac{1}{2}}$ is eighter 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 ∞.
- $a_{n-2} = 0$, $a_n \neq 0$, $a_{n-1} \neq 0$ 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. $\square$

LEMMA 3.4.2. *Let $P(x) = a_0 + a_1x + \cdots + a_nx^n$ be a polynomial over $\mathbb{R}$, satisfying $|P(x)| \le C$ for all $|x| \le D$ ($C, D \in \mathbb{R}_+$). Then*

$$|a_j| \le 8n^{n+1}\frac{C}{D^j} \qquad (j = 0, \dots, n).$$

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. [**36**]: *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

$$t_{n-(2k+1)}^{(n)} = 0 \quad \textit{for} \quad k = 0, \dots, \left[\frac{n-1}{2}\right]$$

and

$$t_{n-(2k)}^{(n)} = (-1)^k \sum_{j=k}^{[\frac{n}{2}]} \binom{n}{2j}\binom{j}{k} \quad \textit{for} \quad k = 0, \dots, \left[\frac{n}{2}\right].$$

The extrema of $T_n(x)$ are given by $\eta_j^{(n)} = \cos\frac{j\pi}{n}$ ($j = 0, \dots, n$). All of them lie in the interval $[-1, 1]$.
Let $C_n = \{P \in \mathcal{P}_n : \max_{j=0,\dots,n} |P(\eta_j^{(n)})| \le 1\}$ and consider $P(x) = a_0 + a_1x + \cdots + a_{n+1}x^{n+1}$. If $n + 1 - j$ is even (or zero) and $P \in C_{n+1}$, then

$$|a_j| \le \left|t_j^{(n+1)}\right|.$$

If $n + 1 - j$ is odd and $P \in C_n$, then

$$|a_j| \le \left|t_j^{(n)}\right|.$$

By assumption, the polynomial

$$\frac{1}{C}P(Dy) = \frac{a_0}{C} + \frac{a_1D^1}{C}y + \cdots + \frac{a_nD^n}{C}y^n$$

belongs to C_{n+1} and C_n. Therefore, since $\binom{p}{q} \le 2^p$,

$$\left|\frac{a_jD^j}{c}\right| \le \max\{\left|t_j^{(n+1)}\right|, \left|t_j^{(n)}\right|\} \le n2^{n+1+[\frac{n+1}{2}]} \le 8n^{n+1} \qquad (j = 0, \dots, n).$$

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

$$N(x) = P(x_0) + P(x_0, x_1)(x - x_0) + \cdots + P(x_0, \dots, x_n)(x - x_0)\cdots(x - x_{n-1})$$

with the divided differences given by

$$\frac{P(x_{j_0}, x_{j_0+1}, \dots, x_{j_0+k-1}) - P(x_{j_0+1}, \dots, x_{j_0+k})}{x_{j_0} - x_{j_0+k}} = P(x_{j_0}, \dots, x_{j_0+k}),$$

for $j_0 = 0, \dots, n-1$. Suppose the nodes are distributed equidistantly. By induction on k we show that

$$|P(x_{j_0}, \dots, x_{j_0+k})| \le \frac{n^k}{k!} \cdot \frac{C}{D^k} \qquad (j_0 = 0, \dots, n).$$

The case $k = 0$ is trivial, since $|P(x_i)| \le C$, for all $i = 0, \dots, n$, by assumption. Let us assume the statement is true for $k-1$, then, since the nodes are distributed equidistantly:

$$\begin{aligned} |P(x_{j_0}, \dots, x_{j_0+k})| &= \frac{|P(x_{j_0}, \dots, x_{j_0+k-1}) - P(x_{j_0+1}, \dots, x_{j_0+k})|}{|x_{j_0} - x_{j_0+k}|} \\ &\le \frac{n}{2kD} \left(|P(x_{j_0}, \dots, x_{j_0+k-1})| + |P(x_{j_0+1}, \dots, x_{j_0+k})|\right) \\ &\le \frac{n}{2kD} \cdot 2 \cdot \frac{n^{k-1}}{(k-1)!} \cdot \frac{C}{D^{k-1}} = \frac{n^k}{k!} \cdot \frac{C}{D^k}. \end{aligned}$$

By expanding $N(x)$, we obtain $N(x) = b_0 + b_1 x + \dots + b_n x^n$ with

$$b_j = \sum_{k=j}^{n} (-1)^{k-j} P(x_0, \dots, x_k) \sum_{0 \le l_1 < \dots < l_{k-j} \le k-1} x_{l_1} \cdots x_{l_{k-j}},$$

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 $n+1$ different nodes is unique, thus, we have $a_j = b_j$ for all $j = 0, \dots, n$, and so

$$\begin{aligned} |a_j| &\le \sum_{k=j}^{n} |P(x_0, \dots, x_k)| \sum_{0 \le l_1 < \dots < l_{k-j} \le k-1} |x_{l_1}| \cdots |x_{l_{k-j}}| \\ &\le \sum_{k=j}^{n} \frac{n^k}{k!} \cdot \frac{C}{D^k} \cdot \binom{k}{k-j} \cdot D^{k-j} \\ &\le \frac{C}{D^j} \sum_{k=j}^{n} n^k \le n^{n+1} \frac{C}{D^j}. \end{aligned}$$

So the proof is complete. $\square$

Lemma 3.4.3. *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.,

$$x_m^n + a_{m,1} x_m^{n-1} + \dots + a_{m,n} = 0 \qquad (m \in \mathbb{N}).$$

Consequently,

$$\begin{aligned} |x_m|^n &= |-a_{m,1} x_m^{n-1} - \dots - a_{m,n}| \\ &\le |a_{m,1}||x_m|^{n-1} + \dots + |a_{m,n}| \qquad (m \in \mathbb{N}). \end{aligned}$$

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 $(x_m)_m$ being unbounded. $\square$

Lemma 3.4.4. *A hyperbolic polynomial*

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

with real coefficients a_i satisfies the following properties:

(1) *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 $y_1 \leq y_2$ (equality means a multiple root) of P' there is a root of P.*
(3) *If $P'(y_0) = 0$ and $P(y_0) \neq 0$, then $P(y_0)P''(y_0) < 0$.*
(4) *If $a_n \neq 0$, then $|a_j| + |a_{j+1}| \neq 0$, for all $j = 1, \dots, n-1$.*
(5) *If $a_{n-1} \neq 0$, then P has a root of the form $x_0 = n\rho\frac{a_n}{a_{n-1}}$ where $0 < \rho \leq 1$, and $P'(x_0)a_{n-1}(-1)^{n-1} \geq 0$.*
(6)

$$a_n a_{n-2} \leq \sum_{j=0}^{n-1} |a_j| \left(n \left| \frac{a_{n-1}}{a_{n-2}} \right| \right)^{n-j} |a_{n-2}| \qquad (a_0 = 1).$$

(7)

$$\begin{aligned} |a_n| \quad \leq \quad & \sum_{j=0}^{n-1} |a_j| \left(2n^2 \left(\min \left(\left| \frac{a_{n-1}}{a_{n-2}} \right|, \left| \frac{a_{n-1}}{a_{n-3}} \right|^{\frac{1}{2}} \right) \right.\right. \\ & \left.\left. + \min \left(\left| \frac{a_{n-2}}{a_{n-3}} \right|, \left| \frac{a_{n-2}}{a_{n-4}} \right|^{\frac{1}{2}} \right) \right) \right)^{n-j} \qquad (a_0 = 1). \end{aligned}$$

PROOF. (1) and (2) are immediate corollaries of Rolle's Theorem which states that the derivative f' of a 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 $i \in \{3, \dots, n\}$ 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 $\left[0, n\frac{a_n}{a_{n-1}}\right]$. It is not possible that all roots are negative, since a_n and a_{n-1} have the same sign. So there are roots $x_{j_1}, \dots, x_{j_k} > n\frac{a_n}{a_{n-1}}$. For a fixed $i \in \{1, \dots, k\}$ we have

$$x_{j_i} > n\frac{a_n}{a_{n-1}} = n\frac{x_1 \cdots x_n}{x_2 \cdots x_n + \cdots + x_1 \cdots x_{n-1}}$$

leading to

$$\frac{x_{j_i}}{x_1} + \cdots + \frac{x_{j_i}}{x_n} > n.$$

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 we can conclude that $x_{j_i} > x_{j_l}$ for one $l \in \{1, \dots, k\} \backslash \{i\}$. And since i was arbitrary, it leads to a contradiction. Therefore the existence follows.
From all such roots choose one of minimal absolute-value x_0. Then $x_0 = 0$, or it

means that P has the same sign inside the segment with endpoints 0 and x_0. In both cases we find that

$$0 \geq P'(x_0)x_0P(0) = P'(x_0)x_0(-1)^n a_n = (-1)^n P'(x_0)n\rho\frac{a_n^2}{a_{n-1}}$$

which is equivalent to $P'(x_0)a_{n-1}(-1)^{n-1} \geq 0$.

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

$$P'(x) = nx^{n-1} - (n-1)a_1x^{n-2} + \cdots + (-1)^{n-2}2a_{n-2}x + (-1)^{n-1}a_{n-1}.$$

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 $y_0^n - a_1y_0^{n-1} + \cdots + (-1)^n a_n = 0$, implying

$$\begin{aligned} |a_n| &\leq |y_0|^n + |a_1||y_0|^{n-1} + \cdots + |a_{n-1}||y_0| \\ &= \sum_{j=0}^{n-1} |a_j| \left(\frac{n-1}{2}|\rho|\left|\frac{a_{n-1}}{a_{n-2}}\right|\right)^{n-j} \\ &\leq \sum_{j=0}^{n-1} |a_j| \left(n\left|\frac{a_{n-1}}{a_{n-2}}\right|\right)^{n-j} \end{aligned}$$

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

$$0 > P(y_0) = y_0^n - a_1y_0^{n-1} + \cdots + (-1)^n a_n$$

by $(-1)^{n-2}a_{n-2}$ gives

$$\begin{aligned} a_na_{n-2} &< (-1)^{n-2}a_{n-2}\left(-y_0^n + a_1y_0^{n-1} + \cdots + (-1)^n a_{n-1}\right) \\ &\leq \sum_{j=0}^{n-1} |a_j| \left(n\left|\frac{a_{n-1}}{a_{n-2}}\right|\right)^{n-j} |a_{n-2}|. \end{aligned}$$

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, i.e.,

$$x^n - a_1x^{n-1} + \cdot + (-1)^n a_n = 0,$$

follows

$$|a_n| \leq \sum_{j=0}^{n-1} |a_j||x|^{n-j}.$$

Let us apply Lemma 3.4.1 to the polynomial P'. Suppose the roots y_i of P' satisfy the inequalities $|y_1| \leq |y_2| \leq \cdots \leq |y_{n-1}|$. Then

$$\begin{aligned} |y_2| &\leq 2(n-1)^2\left(\min\left(\left|\frac{a_{n-1}}{2a_{n-2}}\right|, \left|\frac{a_{n-1}}{3a_{n-3}}\right|^{\frac{1}{2}}\right) + \min\left(\left|\frac{2a_{n-2}}{3a_{n-3}}\right|, \left|\frac{2a_{n-2}}{4a_{n-4}}\right|^{\frac{1}{2}}\right)\right) \\ &\leq 2n^2\left(\min\left(\left|\frac{a_{n-1}}{a_{n-2}}\right|, \left|\frac{a_{n-1}}{a_{n-3}}\right|^{\frac{1}{2}}\right) + \min\left(\left|\frac{a_{n-2}}{a_{n-3}}\right|, \left|\frac{a_{n-2}}{a_{n-4}}\right|^{\frac{1}{2}}\right)\right). \end{aligned}$$

Since, by (2), between y_1 and y_2 there is a root x of P with $|x| \leq |y_2|$, the required inequality follows. □

3.5. 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 hyperbolic polynomials. The essential part of this proof is the following lemma.

LEMMA 3.5.1. *We consider the polynomial*

$$P(t)(x) = \sum_{j=0}^{m-r-1} B_j(t)x^{m-j} + \sum_{j=0}^{r} A_j(t)x^{r-j},$$

which is hyperbolic for all $t \in [-1,1]$. All B_i are bounded functions on $[-1,1]$, and 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,

$$\left|\frac{A_r'(0)}{A_{r-1}(0)}\right| \leq \left(\sup_{i,t}|B_i(t)| + \max_{i,t,j\leq i}|A_i^{(j)}(t)| + \max_t|A_0(t)|^{-1} + 2\right)^C. \tag{3.15}$$

REMARKS. (1) Clearly, the lemma remains true, if, instead of 0, we consider an arbitrary point t_0 and replace the assumptions in the obvious way.
(2) 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 $B_0 \equiv 1$. 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 2.3.3, 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 $P_2(t)(x) = x^{m-1} - D_1(t)x^{m-2} + \cdots + (-1)^{m-1}D_{m-1}(t)$ with $C_1, D_1, \ldots, D_{m-1} \in C^m([-1,1])$ and $C_1(0) = 0$. 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 assumtions, the inequality (3.15) 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.

PROOF. 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 $M_0 < M_1 < \cdots < M_r$, sufficiently large, that all estimates to come in this proof are fulfilled. For example, it is possible to set

$$M_0 = \left(m^m\left(\sup_{i,t}|B_i(t)| + \max_{i,t,j\leq i}|A_i^{(j)}(t)| + \max_t|A_0(t)|^{-1} + 4\right)\right)^{2m}$$

and

$$M_i = M_{i-1}^{4(m+4)^2} \qquad (i = 1, \ldots, r).$$

Let I be the set of indices $i \in \{1, \ldots, r-1\}$ satisfying the following system of conditions:

$(I.1)$ $a_i a_{i-1} \neq 0$ and $\left|\frac{a_i}{a_{i-1}}\right| \leq M_i$

$(I.2)$ $\frac{1}{2} \leq \frac{A_i(t)}{a_i} \leq 2$, if $|t| \leq M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|$

$(I.3)$ $M_0^{-1} \leq \frac{A_{i-1}(t)}{a_{i-1}} \leq M_0$, if $|t| \leq M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|$

$(I.4)$ $|A_j(t)|\left|\frac{a_i}{a_{i-1}}\right|^{i-j} \leq M_i|a_i|$, if $|t| \leq M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|$ and $j = 0, \ldots, i+1$.

And let II be the set of indices $i \in \{2, \ldots, r-1\}$ satisfying these conditions:

$(II.1)$ $a_i a_{i-2} \neq 0$ and $\left|\frac{a_i}{a_{i-2}}\right| \leq M_i$

$(II.2)$ $\frac{1}{2} \le \frac{A_i(t)}{a_i} \le 2$, 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 $j = 0, \ldots, i-1$.

Just to shorten notation let us write $(I.4_{j_0})$ and $(II.4_{j_0})$ for the conditions $(I.4)$ and $(II.4)$ with $j = j_0$, respectively. Define $J = I \cup II$. Note that r cannot be in J, since $a_r = 0$, by assumption, contradicting $(I.1)$ and $(II.1)$.

Claim 1. J is not empty. More precisely: $1 \in I$, if $|a_0a_2| \le a_1^2$, and $2 \in II$, if $|a_0a_2| \ge a_1^2$.

For the hyperbolic polynomial

$$\begin{aligned}\left(\frac{\partial}{\partial x}\right)^{r-2} P(t)(x) &= \sum_{j=0}^{m-r-1} \frac{(m-j)!}{(m-r-j+2)!} B_j(t) x^{m-r-j+2} \\ &\quad + \frac{r!}{2!} A_0(t)x^2 + \frac{(r-1)!}{1!} A_1(t)x + (r-2)! A_2(t),\end{aligned}$$

use lemma 3.4.4(6) to 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 $t \in [-1, 1]$,

$$A_2(t)a_0 \le M_0^{\frac{1}{2}} A_2(t)A_0(t) \le M_0 A_1^2(t). \tag{3.16}$$

Now, consider the case that $|a_0a_2| \le a_1^2$. Then, $a_1 \ne 0$, for otherwise $a_2 = a_1 = 0$, since $|a_0a_2| \le a_1^2$, which contradicts lemma 3.4.4(4). We put $t = \pm M_0^{-1}a_1$ into (3.16) and use Taylor's formula :

$$a_2a_0 \pm a_2^{(1)} M_0^{-1} a_1 a_0 + \frac{A_2^{(2)}(\xi)}{2!} M_0^{-2} a_1^2 a_0 \le M_0 \left(a_1 \pm A_1^{(1)}(\eta) M_0^{-1} a_1\right)^2$$

implies

$$\begin{aligned}\pm a_2^{(1)} M_0^{-1} a_1 a_0 &\le M_0 a_1^2 \pm 2A_1^{(1)}(\eta) a_1^2 + (A_1^{(1)}(\eta))^2 M_0^{-1} a_1^2 \\ &\qquad - a_2 a_0 - \frac{A_2^{(2)}(\xi)}{2!} M_0^{-2} a_1^2 a_0 \\ &\le M_0 a_1^2 + 2|A_1^{(1)}(\eta)| a_1^2 + (A_1^{(1)}(\eta))^2 M_0^{-1} a_1^2 \\ &\qquad + |a_2 a_0| + \frac{|A_2^{(2)}(\xi)|}{2!} M_0^{-2} a_1^2 |a_0| \\ &\le (M_0 + 2M_0 + M_0 + 1 + 1) a_1^2 \\ &\le M_0^2 a_1^2,\end{aligned}$$

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

$$\begin{aligned}|A_2(t)a_0| &= \left|a_2a_0 + a_2^{(1)}ta_0 + \frac{A_2^{(2)}(\xi)}{2!}t^2a_0\right| \\ &\le |a_2a_0| + |a_2^{(1)}||t||a_0| + \frac{|A_2^{(2)}(\xi)|}{2!}|t|^2|a_0| \\ &\le a_1^2 + M_0^4|a_1|M_1^{-1}\left|\frac{a_1}{a_0}\right||a_0| + M_0M_1^{-2}\left|\frac{a_1}{a_0}\right|^2|a_0| \\ &\le M_1a_1^2.\end{aligned}$$

Hence, for $i=1$ the conditions of I are satisfied, and $1 \in J$: $(I.1)$ 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 $\frac{|A_1^{(1)}(\xi)|}{|a_1|}|t| \le M_0M_1^{-1}|a_0|^{-1} \le M_0^2M_1^{-1} \le \frac{1}{2}$. $(I.3)$ follows by definition of M_0, and it implies $(I.4_0)$. $(I.2)$ implies $(I.4_1)$, and $(I.4_2)$ was shown above.
In the second case, when $|a_0a_2| \ge a_1^2$, we find as before that $a_2 \ne 0$. Into (3.16) we put $t = \pm M_0^{-1}\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}}$, and we compute:

$$a_2a_0 \pm a_2^{(1)}M_0^{-1}\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}}a_0 + \frac{A_2^{(2)}(\xi)}{2!}M_0^{-2}\left|\frac{a_2}{a_0}\right|a_0 \le M_0\left(a_1 \pm A_1^{(1)}(\eta)M_0^{-1}\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}}\right)^2$$

implies

$$\begin{aligned}\pm a_2^{(1)}M_0^{-1}\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}}a_0 &\le M_0a_1^2 \pm 2A_1^{(1)}(\eta)a_1\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}} + (A_1^{(1)}(\eta))^2M_0^{-1}\left|\frac{a_2}{a_0}\right| \\ &\quad - a_2a_0 - \frac{A_2^{(2)}(\xi)}{2!}M_0^{-2}\left|\frac{a_2}{a_0}\right|a_0 \\ &\le M_0a_1^2 + 2|A_1^{(1)}(\eta)||a_1|\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}} + (A_1^{(1)}(\eta))^2M_0^{-1}\left|\frac{a_2}{a_0}\right| \\ &\quad + |a_2a_0| + \frac{|A_2^{(2)}(\xi)|}{2!}M_0^{-2}\left|\frac{a_2}{a_0}\right||a_0| \\ &\le (M_0^2 + 2M_0 + M_0^2 + M_0 + M_0)|a_2| \\ &\le M_0^3|a_2|,\end{aligned}$$

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

$$\begin{aligned}|A_1(t)| &= |a_1 + A_1^{(1)}(\xi)t| \le |a_1| + |A_1^{(1)}(\xi)||t| \\ &\le |a_0a_2|^{\frac{1}{2}} + M_0M_2^{-1}\left|\frac{a_2}{a_0}\right|^{\frac{1}{2}} = |a_0a_2|^{\frac{1}{2}}\left(1 + M_0M_2^{-1}\frac{1}{|a_0|}\right) \\ &\le |a_0a_2|^{\frac{1}{2}}\left(1 + M_0^2M_2^{-1}\right) \le 2|a_0a_2|^{\frac{1}{2}}.\end{aligned}$$

Thus, the index $i = 2$ satisfies the conditions of II, and $2 \in J$: $(II.1)$ and $(II.2)$ are clear. $(II.3)$ follows by the definition of M_0, and it implies $(II.4_0)$. $(II.4_1)$ has been shown in the last computation. Therefore the proof of claim 1 is completed.

Claim 2. We want to prove that $r - 1 \in J$.

Suppose otherwise. Let i be the largest index belonging to J, and $i < r - 1$. Then $i + 2 \leq r$. We assert the following implications:

I'. If $i \in I$ and $|a_i a_{i+2}| \leq a_{i+1}^2$, then $i + 1 \in I$.
I''. If $i \in I$ and $|a_i a_{i+2}| \geq a_{i+1}^2$, then $i + 1 \in II$.
II'. If $i \in II$ and $|a_i a_{i+2}| \leq a_{i+1}^2$, then $i + 1 \in I$.
II''. If $i \in II$ and $|a_i a_{i+2}| \geq a_{i+1}^2$, then $i + 1 \in II$.

First of all assume that i satisfies the conditions of I, without specifying the subcases I' and I''. For $|t| \leq M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right|$, consider

$$
\begin{aligned}
\left|\sum_{j=0}^{i} \frac{a_{i+1}^{(j)}}{j!} t^j\right| &= \left|A_{i+1}(t) - \frac{A_{i+1}^{(i+1)}(\xi)}{(i+1)!} t^{i+1}\right| \\
&\leq |A_{i+1}(t)| + \frac{|A_{i+1}^{(i+1)}(\xi)|}{(i+1)!} |t|^{i+1} \\
&\overset{(I.4_{i+1})}{\leq} M_i \left|\frac{a_i^2}{a_{i-1}}\right| + M_0 M_i^{-i-1} \left|\frac{a_i}{a_{i-1}}\right|^{i+1} \\
&\overset{(I.4_0)}{\leq} M_i \left|\frac{a_i^2}{a_{i-1}}\right| + M_0^2 M_i^{-i} |a_i| \left|\frac{a_i}{a_{i-1}}\right| \\
&\leq 2M_i \left|\frac{a_i^2}{a_{i-1}}\right|.
\end{aligned}
$$

Applying lemma 3.4.2, gives

$$
|a_{i+1}^{(j)}| \leq M_0 M_i^{j+1} \left|\frac{a_i^{2-j}}{a_{i-1}^{1-j}}\right| \qquad (j = 0, \ldots, i). \tag{3.17}
$$

Consequently, for $|t| \leq M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right|$, we conclude that

$$
\begin{aligned}
|A'_{i+1}(t)| &= \left|a_{i+1}^{(1)} + a_{i+1}^{(2)} t + \cdots + \frac{a_{i+1}^{(i)}}{(i-1)!} t^{i-1} + \frac{A_{i+1}^{(i+1)}(\xi)}{i!} t^i\right| \\
&\leq \sum_{j=0}^{i-1} \frac{|a_{i+1}^{(j+1)}|}{j!} |t|^j + \frac{|A_{i+1}^{(i+1)}(\xi)|}{i!} |t|^i \\
&\leq \sum_{j=0}^{i-1} \frac{1}{j!} M_0 M_i^{j+2} \left|\frac{a_i^{2-(j+1)}}{a_{i-1}^{1-(j+1)}}\right| M_i^{-j} \left|\frac{a_i}{a_{i-1}}\right|^j + M_0 M_i^{-i} \left|\frac{a_i}{a_{i-1}}\right|^i \\
&\overset{(I.4_0)}{\leq} \sum_{j=0}^{i-1} \frac{1}{j!} M_0 M_i^2 |a_i| + M_0^2 M_i^{-i+1} |a_i| \\
&\leq M_0^2 M_i^2 |a_i|.
\end{aligned} \tag{3.18}
$$

For $|t| \leq \frac{1}{2} M_0^{-2} M_i^{-2} \left|\frac{a_{i+1}}{a_i}\right| \left(\overset{(I.4_{i+1})}{\leq} \frac{1}{2} M_0^{-2} M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right| \leq M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right|\right)$, we obtain, by the mean value theorem:

$$\begin{aligned} |A_{i+1}(t) - a_{i+1}| &= |A'_{i+1}(\xi)||t| \\ &\leq M_0^2 M_i^2 |a_i| \cdot \frac{1}{2} M_0^{-2} M_i^{-2} \left|\frac{a_{i+1}}{a_i}\right| \\ &= \frac{1}{2}|a_{i+1}|. \end{aligned} \tag{3.19}$$

Consider the hyperbolic polynomial

$$\begin{aligned} \left(\frac{\partial}{\partial x}\right)^{r-(i+2)} P(t)(x) &= \sum_{j=0}^{m-r-1} \frac{(m-j)!}{(m-r+(i+2)-j)!} B_j(t) x^{m-r+(i+2)-j} \\ &\quad + \frac{r!}{(i+2)!} A_0(t) x^{i+2} + \cdots + (r-(i+2))! A_{i+2}(t), \end{aligned}$$

and use lemma 3.4.4(7) to obtain

$$\begin{aligned} |A_{i+2}(t)| &\leq M_0 \sum_{j=i+2}^{m} \left(\left|\frac{A_i(t)}{A_{i-1}(t)}\right| + \left|\frac{A_{i+1}(t)}{A_i(t)}\right|\right)^j \\ &\quad + M_0 \sum_{j=0}^{i+1} |A_j(t)| \left(\left|\frac{A_i(t)}{A_{i-1}(t)}\right| + \left|\frac{A_{i+1}(t)}{A_i(t)}\right|\right)^{i+2-j}. \end{aligned}$$

The inequalities $(I.2)$, $(I.3)$ and $(I.4_{i+1})$ provide, for $|t| \leq M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right|$, the following estimates:

$$\left|\frac{A_i(t)}{A_{i-1}(t)}\right| \leq 2M_0 \left|\frac{a_i}{a_{i-1}}\right|$$

and

$$\left|\frac{A_{i+1}(t)}{A_i(t)}\right| \leq M_i \left|\frac{a_i}{A_i(t)}\right| \left|\frac{a_i}{a_{i-1}}\right| \leq 2M_i \left|\frac{a_i}{a_{i-1}}\right|. \tag{3.20}$$

Therefore, for $|t| \leq M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right|$,

$$\begin{aligned} |A_{i+2}(t)| &\leq M_0 \sum_{j=i+2}^{m} \left(4M_i \left|\frac{a_i}{a_{i-1}}\right|\right)^j \\ &\quad + M_0 \sum_{j=0}^{i+1} |A_j(t)| \left(4M_i \left|\frac{a_i}{a_{i-1}}\right|\right)^{i+2-j}. \end{aligned}$$

Consider the first sum on the right-hand side:

$$\begin{aligned} \sum_{j=i+2}^{m} \left(4M_i \left|\frac{a_i}{a_{i-1}}\right|\right)^j &= \sum_{j=0}^{m-i-2} (4M_i)^{i+2+j} \left|\frac{a_i}{a_{i-1}}\right|^i \left|\frac{a_i}{a_{i-1}}\right|^{j+2} \\ &\overset{(I.4_0)}{\leq} \sum_{j=0}^{m-i-2} (4M_i)^{i+2+j} M_0 M_i |a_i| \left|\frac{a_i}{a_{i-1}}\right|^{j+2} \\ &\overset{(I.1)}{\leq} M_0 \sum_{j=0}^{m-i-2} (4M_i)^{i+2+j} M_i^{j+1} \left|\frac{a_i^3}{a_{i-1}^2}\right| \\ &\leq M_0 M_i^{2m} \left|\frac{a_i^3}{a_{i-1}^2}\right|. \end{aligned}$$

Thus, for $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|$,

$$\begin{aligned}
|A_{i+2}(t)| &\le M_0^2 M_i^{2m}\left|\frac{a_i^3}{a_{i-1}^2}\right| + M_0\sum_{j=0}^{i+1}|A_j(t)|\left(4M_i\left|\frac{a_i}{a_{i-1}}\right|\right)^{i+2-j} \\
&\overset{(I.4_j)}{\le} M_0^2 M_i^{2m}\left|\frac{a_i^3}{a_{i-1}^2}\right| + M_0\sum_{j=0}^{i+1}M_i|a_i|\left|\frac{a_i}{a_{i-1}}\right|^{j-i}\left(4M_i\left|\frac{a_i}{a_{i-1}}\right|\right)^{i+2-j} \\
&\le M_i^{2m+2}\left|\frac{a_i^3}{a_{i-1}^2}\right|.
\end{aligned}$$

In the same way as above we conclude, for $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|$, that

$$\begin{aligned}
\left|\sum_{j=0}^{i+1}\frac{a_{i+2}^{(j)}}{j!}t^j\right| &= \left|A_{i+2}(t) - \frac{A_{i+2}^{(i+2)}(\xi)}{(i+2)!}t^{i+2}\right| \\
&\le |A_{i+2}(t)| + \frac{|A_{i+2}^{(i+2)}(\xi)|}{(i+2)!}|t|^{i+2} \\
&\le M_i^{2m+2}\left|\frac{a_i^3}{a_{i-1}^2}\right| + M_0 M_i^{-i-2}\left|\frac{a_i}{a_{i-1}}\right|^{i+2} \\
&\overset{(I.4_0)}{\le} M_i^{2m+2}\left|\frac{a_i^3}{a_{i-1}^2}\right| + M_0^2 M_i^{-i-1}|a_i|\left|\frac{a_i}{a_{i-1}}\right|^2 \\
&\le 2M_i^{2m+2}\left|\frac{a_i^3}{a_{i-1}^2}\right|.
\end{aligned}$$

Use again lemma 3.4.2 to obtain:

$$|a_{i+2}^{(j)}| \le M_0 M_i^{2m+2+j}\left|\frac{a_i^{3-j}}{a_{i-1}^{2-j}}\right| \qquad (j = 0, \ldots, i+1). \tag{3.21}$$

We apply this result to estimate:

$$\begin{aligned}
|A_{i+2}''(t)| &= \left|a_{i+2}^{(2)} + a_{i+2}^{(3)}t + \cdots + \frac{a_{i+2}^{(i+1)}}{(i-1)!}t^{i-1} + \frac{A_{i+2}^{(i+2)}(\xi)}{i!}t^i\right| \\
&\le \sum_{j=0}^{i-1}\left|a_{i+2}^{(j+2)}\right||t|^j + M_0|t|^i \\
&\le \sum_{j=0}^{i-1}M_0 M_i^{2m+2+(j+2)}\left|\frac{a_i^{3-(j+2)}}{a_{i-1}^{2-(j+2)}}\right| M_i^{-j}\left|\frac{a_i}{a_{i-1}}\right|^j \\
&\quad + M_0 M_i^{-i}\left|\frac{a_i}{a_{i-1}}\right|^i \\
&\overset{(I.4_0)}{\le} \sum_{j=0}^{i-1}M_0 M_i^{2m+4}|a_i| + M_0^2 M_i^{-i+1}|a_i| \\
&\le M_i^{2m+6}|a_i|,
\end{aligned} \tag{3.22}$$

for $|t| \le M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|$.

Consider again the polynomial $\left(\frac{\partial}{\partial x}\right)^{r-(i+2)} P(t)(x)$ and apply lemma 3.4.4(6):

$$A_{i+2}(t)A_i(t) \leq \frac{1}{4}M_0 \left(\sum_{j=i+2}^{m} \left| \frac{A_{i+1}(t)}{A_i(t)} \right|^j + \sum_{j=0}^{i+1} |A_j(t)| \left| \frac{A_{i+1}(t)}{A_i(t)} \right|^{i+2-j} \right) |A_i(t)|.$$

By $(I.2)$ and (3.20), we find that, for $|t| \leq M_i^{-1} \left| \frac{a_i}{a_{i-1}} \right|$,

$$\begin{aligned} A_{i+2}(t)a_i \quad &\leq \quad M_0 \left(\sum_{j=i+2}^{m} \left| \frac{A_{i+1}(t)}{A_i(t)} \right|^j + \sum_{j=0}^{i+1} |A_j(t)| \left| \frac{A_{i+1}(t)}{A_i(t)} \right|^{i+2-j} \right) |a_i| \\ &\leq \quad M_0 \left| \frac{A_{i+1}(t)}{A_i(t)} \right|^2 \left(\sum_{j=i}^{m-2} \left(2M_i \left| \frac{a_i}{a_{i-1}} \right| \right)^j \right. \\ &\qquad \left. + \sum_{j=0}^{i+1} |A_j(t)| \left(2M_i \left| \frac{a_i}{a_{i-1}} \right| \right)^{i-j} \right) |a_i|. \end{aligned}$$

Consider the first sum on the right-hand side:

$$\begin{aligned} \sum_{j=i}^{m-2} (2M_i)^j \left| \frac{a_i}{a_{i-1}} \right|^j \quad &= \quad \sum_{j=0}^{m-i-2} (2M_i)^{i+j} \left| \frac{a_i}{a_{i-1}} \right|^i \left| \frac{a_i}{a_{i-1}} \right|^j \\ &\overset{(I.4_0)}{\leq} \quad \sum_{j=0}^{m-i-2} (2M_i)^{i+j} M_0 M_i |a_i| \left| \frac{a_i}{a_{i-1}} \right|^j \\ &\overset{(I.1)}{\leq} \quad \sum_{j=0}^{m-i-2} (2M_i)^{i+j} M_0 M_i^{j+1} |a_i| \\ &\leq \quad M_i^{2m} |a_i|. \end{aligned}$$

Consequently, using $(I.4_j)$, we obtain, for $|t| \leq M_i^{-1} \left| \frac{a_i}{a_{i-1}} \right|$,

$$\begin{aligned} A_{i+2}(t)a_i \quad &\leq \quad M_0 \left| \frac{A_{i+1}(t)}{A_i(t)} \right|^2 \left(M_i^{2m} |a_i| \right. \\ &\qquad \left. + \sum_{j=0}^{i+1} M_i |a_i| \left| \frac{a_i}{a_{i-1}} \right|^{j-i} (2M_i)^{i-j} \left| \frac{a_i}{a_{i-1}} \right|^{i-j} \right) |a_i| \\ &\leq \quad 2M_0 M_i^{2m} |A_{i+1}(t)|^2 \left| \frac{a_i}{A_i(t)} \right|^2 \\ &\leq \quad M_i^{2m+1} |A_{i+1}(t)|^2. \end{aligned} \tag{3.23}$$

All we have done till now is true in the case that $i \in I$. In the following we want to consider seperately the subcases I' and I''.

In the subcase I' we have $|a_i a_{i+2}| \leq a_{i+1}^2$. Then $a_{i+1} \neq 0$ (otherwise $a_{i+1} = a_{i+2} = 0$, since $(I.1)$ tells us that $a_i \neq 0$, contradicting lemma 3.4.4(4)). Inequality (3.19) implies, for $|t| \leq \frac{1}{2} M_0^{-2} M_i^{-2} \left| \frac{a_{i+1}}{a_i} \right|$, that

$$-\frac{1}{2}|a_{i+1}| \leq A_{i+1}(t) - a_{i+1} \leq \frac{1}{2}|a_{i+1}|,$$

whence

$$\frac{1}{2} \leq \frac{A_{i+1}(t)}{a_{i+1}} \leq 2.$$

Thus, $(I.2)$ is satisfied for the index $i+1$. Setting $t = \pm M_i^{-3}\left|\frac{a_{i+1}}{a_i}\right|$ $\left(\overset{(I.4_{i+1})}{\leq} M_i^{-2}\left|\frac{a_i}{a_{i-1}}\right| \leq M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|\right)$ into (3.23), remembering (3.22) and using the previous result, we conclude as follows:

$$a_i a_{i+2} \pm a_i a_{i+2}^{(1)} M_i^{-3}\left|\frac{a_{i+1}}{a_i}\right| + a_i \frac{A_{i+2}^{(2)}(\xi)}{2!} M_i^{-6}\left|\frac{a_{i+1}}{a_i}\right|^2 \leq 4M_i^{2m+1}|a_{i+1}|^2$$

implies

$$\begin{aligned}
\pm a_i a_{i+2}^{(1)} M_i^{-3}\left|\frac{a_{i+1}}{a_i}\right| &\leq 4M_i^{2m+1}|a_{i+1}|^2 - a_i a_{i+2} \\
&\quad - a_i \frac{A_{i+2}^{(2)}(\xi)}{2!} M_i^{-6}\left|\frac{a_{i+1}}{a_i}\right|^2 \\
&\leq 4M_i^{2m+1}|a_{i+1}|^2 + |a_i a_{i+2}| \\
&\quad + |a_i|\frac{|A_{i+2}^{(2)}(\xi)|}{2!} M_i^{-6}\left|\frac{a_{i+1}}{a_i}\right|^2 \\
&\overset{(3.22)}{\leq} 4M_i^{2m+1}|a_{i+1}|^2 + |a_{i+1}|^2 + M_i^{2m}|a_{i+1}|^2 \\
&\leq M_i^{2m+2}|a_{i+1}|^2,
\end{aligned}$$

whence

$$|a_{i+2}^{(1)}| \leq M_i^{2m+5}|a_{i+1}|.$$

Thus, we have, for $|t| \leq M_{i+1}^{-1}\left|\frac{a_{i+1}}{a_i}\right|$ $\left(\leq M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right| \text{ as seen before}\right)$,

$$\begin{aligned}
|A_{i+2}(t)a_i| &= \left|a_i a_{i+2} + a_i a_{i+2}^{(1)} t + a_i \frac{A_{i+2}^{(2)}(\xi)}{2!} t^2\right| \\
&\leq |a_i a_{i+2}| + |a_i||a_{i+2}^{(1)}||t| + |a_i|\frac{|A_{i+2}^{(2)}(\xi)|}{2!}|t|^2 \\
&\overset{(3.22)}{\leq} |a_{i+1}|^2 + |a_i| M_i^{2m+5}|a_{i+1}| M_{i+1}^{-1}\left|\frac{a_{i+1}}{a_i}\right| \\
&\quad + |a_i| M_i^{2m+6}|a_i| M_{i+1}^{-2}\left|\frac{a_{i+1}}{a_i}\right|^2 \\
&\leq 2|a_{i+1}|^2.
\end{aligned} \tag{3.24}$$

For $|t| \leq M_{i+1}^{-1}\left|\frac{a_{i+1}}{a_i}\right|$ and $j = 0, \ldots, i+1$, consider

$$\begin{aligned}
|A_j(t)|\left|\frac{a_{i+1}}{a_i}\right|^{i+1-j} &= |A_j(t)|\left|\frac{a_{i+1}}{a_i}\right|\left|\frac{a_{i+1}}{a_i}\right|^{i-j} \\
&\overset{(I.4_{i+1})}{\leq} |A_j(t)|\left|\frac{a_{i+1}}{a_i}\right| M_i^{i-j}\left|\frac{a_i}{a_{i-1}}\right|^{i-j} \\
&\overset{(I.4_j)}{\leq} M_i^{i+1-j}|a_{i+1}| \\
&\leq M_{i+1}|a_{i+1}|.
\end{aligned}$$

Now we are able to see that the index $i+1$ satisfies the conditions of I: $(I.1)$ is clear. We have already seen $(I.2)$. $(I.3)$ is true since $i \in I$. $(I.4_j)$, for $j = 0, \ldots, i+1$, has been shown in the previous computation, and $(I.4_{i+2})$ corresponds to (3.24).

Consider now the subcase I'', where $|a_i a_{i+2}| \geq a_{i+1}^2$. Then, by lemma 3.4.4(4), $a_{i+2} \neq 0$.
Setting $t = \pm M_i^{-2m-3} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} \left(\overset{(3.21_0)}{\leq} M_i^{-m-2} M_0^{\frac{1}{2}} \left|\frac{a_i}{a_{i-1}}\right| \leq M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right|\right)$ into (3.23) gives

$$\begin{aligned} a_i a_{i+2} \pm a_i a_{i+2}^{(1)} M_i^{-2m-3} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} + a_i \frac{A_{i+2}^{(2)}(\xi)}{2!} M_i^{-4m-6} \left|\frac{a_{i+2}}{a_i}\right| \\ \leq \quad M_i^{2m+1} \left| a_{i+1} \pm A_{i+1}^{(1)}(\eta) M_i^{-2m-3} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} \right|^2 . \end{aligned}$$

This implies, by using (3.18) and (3.22),

$$\begin{aligned} & \pm a_i a_{i+2}^{(1)} M_i^{-2m-3} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} \\ \leq \quad & M_i^{2m+1} a_{i+1}^2 \pm 2 a_{i+1} A_{i+1}^{(1)}(\eta) M_i^{-2} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} + (A_{i+1}^{(1)}(\eta))^2 M_i^{-2m-5} \left|\frac{a_{i+2}}{a_i}\right| \\ & - a_i a_{i+2} - a_i \frac{A_{i+2}^{(2)}(\xi)}{2!} M_i^{-4m-6} \left|\frac{a_{i+2}}{a_i}\right| \\ \leq \quad & M_i^{2m+1} a_{i+1}^2 + 2|a_{i+1}||A_{i+1}^{(1)}(\eta)| M_i^{-2} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} + (A_{i+1}^{(1)}(\eta))^2 M_i^{-2m-5} \left|\frac{a_{i+2}}{a_i}\right| \\ & + |a_i a_{i+2}| + |a_i| \frac{|A_{i+2}^{(2)}(\xi)|}{2!} M_i^{-4m-6} \left|\frac{a_{i+2}}{a_i}\right| \\ \leq \quad & \left(M_i^{2m+1} + 2M_0^2 + M_0^4 M_i^{-2m-1} + 1 + \frac{1}{2} M_i^{-2m} \right) |a_i a_{i+2}| \\ \leq \quad & 5 M_i^{2m+1}. \end{aligned}$$

So we get

$$|a_{i+2}^{(1)}| \leq 5 M_i^{4m+4} |a_i a_{i+2}|^{\frac{1}{2}}.$$

For $|t| \leq M_i^{-4m-6} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}}$, this gives:

$$\begin{aligned} |A_{i+2}(t) - a_{i+2}| \quad &= \quad \left| a_{i+2}^{(1)} t + \frac{A_{i+2}^{(2)}(\xi)}{2!} t^2 \right| \leq |a_{i+2}^{(1)}||t| + \frac{|A_{i+2}^{(2)}(\xi)|}{2!} |t|^2 \\ &\overset{(3.22)}{\leq} \quad 5 M_i^{-2} |a_{i+2}| + \frac{1}{2} M_i^{-6m-6} |a_{i+2}| \leq \frac{1}{2} |a_{i+2}|, \end{aligned}$$

whence $i+2$ satisfies $(II.2)$.
Finally, consider

$$\begin{aligned} |A_j(t)| \left|\frac{a_{i+2}}{a_i}\right|^{\frac{i+2-j}{2}} \quad &= \quad |A_j(t)| \left|\frac{a_{i+2}}{a_i}\right| \left|\frac{a_{i+2}}{a_i}\right|^{\frac{i-j}{2}} \\ &\overset{(3.21_0)}{\leq} \quad |A_j(t)| \left|\frac{a_{i+2}}{a_i}\right| \left(M_0 M_i^{2m+2} \left|\frac{a_i}{a_{i-1}}\right|^2 \right)^{\frac{i-j}{2}} \\ &\overset{(I.4_j)}{\leq} \quad M_i |a_i| \left|\frac{a_i}{a_{i-1}}\right|^{j-i} \left|\frac{a_{i+2}}{a_i}\right| M_0^{\frac{i-j}{2}} M_i^{(m+1)(i-j)} \left|\frac{a_i}{a_{i-1}}\right|^{i-j} \\ &\leq \quad M_i^{(m+1)^2} |a_{i+2}| \end{aligned}$$

which works for $|t| \leq M_i^{-1} \left|\frac{a_i}{a_{i-1}}\right|$ and $j = 0, \ldots, i+1$.

Thus, the index $i+2$ satisfies the conditions of II: $(II.1)$ and $(II.2)$ are clear. $(II.3)$ is true, since $i \in I$, and $(II.4)$ has been shown in the last estimate.

Let us investigate now the case that i belongs to II. We use lemma 3.4.4(7) for the hyperbolic polynomial

$$\begin{aligned}\left(\frac{\partial}{\partial x}\right)^{r-(i+1)} P(t)(x) &= \sum_{j=0}^{m-r-1} \frac{(m-j)!}{(m-r+(i+1)-j)!} B_j(t) x^{m-r+(i+1)-j} \\ &\quad + \frac{r!}{(i+1)!} A_0(t) x^{i+1} + \cdots + (r-(i+1))! A_{i+1}(t)\end{aligned}$$

and obtain

$$\begin{aligned}|A_{i+1}(t)| &\leq M_0 \sum_{j=i+1}^{m} \left(\left| \frac{A_{i-1}(t)}{A_{i-2}(t)} \right| + \left| \frac{A_i(t)}{A_{i-2}(t)} \right|^{\frac{1}{2}} \right)^j \\ &\quad + M_0 \sum_{j=0}^{i} |A_j(t)| \left(\left| \frac{A_{i-1}(t)}{A_{i-2}(t)} \right| + \left| \frac{A_i(t)}{A_{i-2}(t)} \right|^{\frac{1}{2}} \right)^{i+1-j}.\end{aligned}$$

For $|t| \leq M_i^{-1} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{1}{2}}$, we find following estimates:

$$\left| \frac{A_{i-1}(t)}{A_{i-2}(t)} \right| \overset{(II.4_{i-1})}{\leq} \frac{M_i |a_i a_{i-2}|^{\frac{1}{2}}}{|A_{i-2}(t)|} \overset{(II.3)}{\leq} M_0 M_i \left| \frac{a_i}{a_{i-2}} \right|^{\frac{1}{2}}$$

and

$$\left| \frac{A_i(t)}{A_{i-2}(t)} \right|^{\frac{1}{2}} \overset{(II.2),(II.3)}{\leq} (2M_0)^{\frac{1}{2}} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{1}{2}}. \tag{3.25}$$

Use them to estimate the first sum on the right-hand side:

$$\begin{aligned}\sum_{j=i+1}^{m} \left(\left| \frac{A_{i-1}(t)}{A_{i-2}(t)} \right| + \left| \frac{A_i(t)}{A_{i-2}(t)} \right|^{\frac{1}{2}} \right)^j &\leq \sum_{j=i+1}^{m} (M_0^2 M_i)^j \left| \frac{a_i}{a_{i-2}} \right|^{\frac{j}{2}} \\ &= \sum_{j=1}^{m-i} (M_0^2 M_i)^{i+j} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{i}{2}} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{j}{2}} \\ &\overset{(II.4_0)}{\leq} \sum_{j=1}^{m-i} (M_0^2 M_i)^{i+j} M_0 M_i |a_i| \left| \frac{a_i}{a_{i-2}} \right|^{\frac{j}{2}} \\ &= M_0 M_i |a_i| \left| \frac{a_i}{a_{i-2}} \right|^{\frac{1}{2}} \sum_{j=1}^{m-i} (M_0^2 M_i)^{i+j} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{j-1}{2}} \\ &\overset{(II.1)}{\leq} M_0 M_i |a_i| \left| \frac{a_i}{a_{i-2}} \right|^{\frac{1}{2}} \sum_{j=1}^{m-i} (M_0^2 M_i)^{i+j} M_i^{\frac{j-1}{2}} \\ &\leq M_0^{2m+2} M_i^{m+1} |a_i| \left| \frac{a_i}{a_{i-2}} \right|^{\frac{1}{2}}.\end{aligned}$$

Therefore, we get, for $|t| \leq M_i^{-1} \left| \frac{a_i}{a_{i-2}} \right|^{\frac{1}{2}}$,

$$\begin{aligned}
|A_{i+1}(t)| \quad & \leq \quad M_0^{2m+3} M_i^{m+1} |a_i| \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \\
& \quad + M_0 \sum_{j=0}^{i-1} |A_j(t)| (M_0^2 M_i)^{i+1-j} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{i+1-j}{2}} \\
& \quad + M_0 |A_i(t)| M_0^2 M_i \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \\
& \overset{(II.2),(II.4_j)}{\leq} \quad M_0^{2m+3} M_i^{m+1} |a_i| \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \\
& \quad + M_0 M_i |a_i| \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \sum_{j=0}^{i-1} (M_0^2 M_i)^{i+1-j} \\
& \quad + 2 M_0^3 M_i |a_i| \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \\
& \leq \quad 3 M_0^{2m+3} M_i^{m+1} |a_i| \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \\
& \leq \quad M_i^{m+2} |a_i| \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}. \qquad (3.26)
\end{aligned}$$

Next we use lemma 3.4.4(7) for $\left(\frac{\partial}{\partial x}\right)^{r-(i+2)} P(t)(x)$, and we find

$$\begin{aligned}
|A_{i+2}(t)| \quad \leq \quad & M_0 \sum_{j=i+2}^{m} \left(\left|\frac{A_{i+1}(t)}{A_i(t)}\right| + \left|\frac{A_i(t)}{A_{i-2}(t)}\right|^{\frac{1}{2}} \right)^{j} \\
& + M_0 \sum_{j=0}^{i+1} |A_j(t)| \left(\left|\frac{A_{i+1}(t)}{A_i(t)}\right| + \left|\frac{A_i(t)}{A_{i-2}(t)}\right|^{\frac{1}{2}} \right)^{i+2-j}.
\end{aligned}$$

Considering (3.25) and the fact that, for $|t| \leq M_i^{-1} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$,

$$\left|\frac{A_{i+1}(t)}{A_i(t)}\right| \overset{(3.26)}{\leq} M_i^{m+2} \left|\frac{a_i}{A_i(t)}\right| \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \overset{(II.2)}{\leq} 2 M_i^{m+2} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}, \qquad (3.27)$$

we estimate the first sum on the right-hand side

$$\begin{aligned}
\sum_{j=i+2}^{m} \left(\left|\frac{A_{i+1}(t)}{A_i(t)}\right| + \left|\frac{A_i(t)}{A_{i-2}(t)}\right|^{\frac{1}{2}} \right)^{j} & \leq \sum_{j=i+2}^{m} (M_0 M_i^{m+2})^j \left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}} \\
& = \sum_{j=2}^{m-i} (M_0 M_i^{m+2})^{i+j} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{i}{2}} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}} \\
& \overset{(II.4_0)}{\leq} \sum_{j=2}^{m-i} (M_0 M_i^{m+2})^{i+j} M_0 M_i |a_i| \left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}} \\
& = M_0 M_i |a_i| \left|\frac{a_i}{a_{i-2}}\right| \sum_{j=2}^{m-i} (M_0 M_i^{m+2})^{i+j} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{j-2}{2}}
\end{aligned}$$

$$
\begin{aligned}
&\overset{(II.1)}{\leq} && M_0 M_i |a_i| \left|\frac{a_i}{a_{i-2}}\right| \sum_{j=2}^{m-i} (M_0 M_i^{m+2})^{i+j} M_i^{\frac{j-2}{2}} \\
&\leq && M_0^{m+2} M_i^{(m+1)^2} |a_i| \left|\frac{a_i}{a_{i-2}}\right|.
\end{aligned}
$$

So, for $|t| \leq M_i^{-1} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$, this gives:

$$
\begin{aligned}
|A_{i+2}(t)| &\leq && M_0^{m+3} M_i^{(m+1)^2} |a_i| \left|\frac{a_i}{a_{i-2}}\right| \\
& && + M_0 \sum_{j=0}^{i-1} |A_j(t)| (M_0 M_i^{m+2})^{i+2-j} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{i+2-j}{2}} \\
& && + M_0 |A_i(t)| (M_0 M_i^{m+2})^2 \left|\frac{a_i}{a_{i-2}}\right| \\
& && + M_0 |A_{i+1}(t)| M_0 M_i^{m+2} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \\
&\overset{(II.2),(II.4_j),(3.26)}{\leq} && M_0^{m+3} M_i^{(m+1)^2} |a_i| \left|\frac{a_i}{a_{i-2}}\right| \\
& && + M_0 M_i |a_i| \left|\frac{a_i}{a_{i-2}}\right| \sum_{j=0}^{i-1} (M_0 M_i^{m+2})^{i+2-j} \\
& && + 2 M_0^3 M_i^{2m+4} |a_i| \left|\frac{a_i}{a_{i-2}}\right| \\
& && + M_0^2 M_i^{2m+4} |a_i| \left|\frac{a_i}{a_{i-2}}\right| \\
&\leq && M_i^{(m+1)^2+1} |a_i| \left|\frac{a_i}{a_{i-2}}\right|.
\end{aligned}
\tag{3.28}
$$

From inequalities (3.26) and (3.28) we can derive the following estimates, just applying lemma 3.4.2 as we did before:

$$
|a_{i+1}^{(j)}| \leq M_0 M_i^{m+2+j} |a_i| \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1-j}{2}} \qquad (j = 0, \dots, i) \tag{3.29}
$$

and

$$
|a_{i+2}^{(j)}| \leq M_0 M_i^{(m+1)^2+1+j} |a_i| \left|\frac{a_i}{a_{i-2}}\right|^{\frac{2-j}{2}} \qquad (j = 0, \dots, i+1). \tag{3.30}
$$

For $|t| \leq M_i^{-1} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$, this yields

$$
\begin{aligned}
|A'_{i+1}(t)| &= \left| a_{i+1}^{(1)} + a_{i+1}^{(2)} t + \cdots + \frac{a_{i+1}^{(i)}}{(i-1)!} t^{i-1} + \frac{A_{i+1}^{(i+1)}(\xi)}{i!} t^i \right| \\
&\leq \sum_{j=0}^{i-1} |a_{i+1}^{(j+1)}| |t|^j + M_0 |t|^i
\end{aligned}
$$

$$
\begin{aligned}
&\overset{(3.29)}{\leq} \sum_{j=0}^{i-1} M_0 M_i^{m+2+j+1}|a_i| \left|\frac{a_i}{a_{i-2}}\right|^{-\frac{j}{2}} M_i^{-j} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}} \\
&\quad + M_0 M_i^{-i} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{i}{2}} \\
&\overset{(II.4_0)}{\leq} \sum_{j=0}^{i-1} M_0 M_i^{m+3}|a_i| + M_0^2 M_i^{-i+1}|a_i| \\
&\leq \frac{1}{2} M_i^{m+4}|a_i|. \qquad (3.31)
\end{aligned}
$$

For $|t| \leq M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$, we consider

$$
\begin{aligned}
|A''_{i+2}(t)| &= \left|a_{i+2}^{(2)} + a_{i+2}^{(3)} t + \cdots + \frac{a_{i+2}^{(i+1)}}{(i-1)!} t^{i-1} + \frac{A_{i+2}^{(i+2)}(\xi)}{i!} t^i \right| \\
&\leq \sum_{j=0}^{i-1} |a_{i+2}^{(j+2)}||t|^j + M_0 |t|^i \\
&\overset{(3.30)}{\leq} \sum_{j=0}^{i-1} M_0 M_i^{(m+1)^2+1+j+2}|a_i| \left|\frac{a_i}{a_{i-2}}\right|^{-\frac{j}{2}} M_i^{-j} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}} \\
&\quad + M_0 M_i^{-i} \left|\frac{a_i}{a_{i-2}}\right|^{\frac{i}{2}} \\
&\overset{(II.4_0)}{\leq} \sum_{j=0}^{i-1} M_0 M_i^{(m+1)^2+3}|a_i| + M_0^2 M_i^{-i+1}|a_i| \\
&\leq M_i^{(m+1)^2+4}|a_i|. \qquad (3.32)
\end{aligned}
$$

Applying lemma 3.4.4(6) to $\left(\frac{\partial}{\partial x}\right)^{r-(i+2)} P(t)(x)$ and using $(II.2)$, gives:

$$
A_{i+2}(t)a_i \leq M_0 \left(\sum_{j=i+2}^{m} \left|\frac{A_{i+1}(t)}{A_i(t)}\right|^j + \sum_{j=0}^{i+1} |A_j(t)| \left|\frac{A_{i+1}(t)}{A_i(t)}\right|^{i+2-j} \right) |a_i|.
$$

By (3.27), we have, for $|t| \leq M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$,

$$
\begin{aligned}
A_{i+2}(t)a_i \leq{} & M_0 \left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2 \sum_{j=i}^{m-2} \left(2M_i^{m+2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\right)^j |a_i| \\
& + M_0 \left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2 \sum_{j=0}^{i+1} |A_j(t)| \left(2M_i^{m+2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\right)^{i-j} |a_i|.
\end{aligned}
$$

Consider the first sum on the right-hand side:

$$\begin{aligned}
\sum_{j=i}^{m-2}\left(2M_i^{m+2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\right)^j &= \sum_{j=0}^{m-i-2}(2M_i^{m+2})^{i+j}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{i}{2}}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}}\\
&\overset{(II.4_0)}{\leq} \sum_{j=0}^{m-i-2}(2M_i^{m+2})^{i+j}M_0M_i|a_i|\left|\frac{a_i}{a_{i-2}}\right|^{\frac{j}{2}}\\
&\overset{(II.1)}{\leq} \sum_{j=0}^{m-i-2}(2M_i^{m+2})^{i+j}M_0M_iM_i^{\frac{j}{2}}|a_i|\\
&\leq M_i^{m^2+m}|a_i|.
\end{aligned}$$

Therefore, for $|t|\leq M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$,

$$\begin{aligned}
A_{i+2}(t)a_i &\overset{(II.4_j)}{\leq}\\
&M_0M_i^{m^2+m}\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2|a_i|^2\\
&+M_0\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2|a_i|\sum_{j=0}^{i-1}M_i|a_i|\left|\frac{a_i}{a_{i-2}}\right|^{\frac{j-i}{2}}\left(2M_i^{m+2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\right)^{i-j}\\
&+M_0\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2|a_i||A_i(t)|\\
&+M_0\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2|a_i||A_{i+1}(t)|\left(2M_i^{m+2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\right)^{-1}\\
&\overset{(II.2),(3.26)}{\leq} M_0M_i^{m^2+m}\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2|a_i|^2+M_0^2M_i^{m^2+m}\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2|a_i|^2\\
&+2M_0\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2|a_i|^2+\frac{1}{2}M_0\left|\frac{A_{i+1}(t)}{A_i(t)}\right|^2|a_i|^2\\
&\overset{(II.2)}{\leq} M_i^{(m+1)^2}|A_{i+1}(t)|^2. \qquad (3.33)
\end{aligned}$$

As we did before, let us specify now the subcases II' and II''.

If i belongs to II', we have $|a_ia_{i+2}|\leq a_{i+1}^2$. Note the $a_{i+1}\neq 0$ (otherwise lemma 3.4.4(4) is harmed).

For $|t|\leq M_i^{-m-4}\left|\frac{a_{i+1}}{a_i}\right|\ \left(\overset{(3.29_0)}{\leq} M_0M_i^{-2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\leq M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\right)$, we find, by (3.31), that

$$|A_{i+1}(t)-a_{i+1}|=|A'_{i+1}(\xi)||t|\leq\frac{1}{2}|a_{i+1}|, \qquad (3.34)$$

which shows that $i+1$ satisfies $(I.2)$.

Put $t=\pm M_i^{-2(m+1)^2}\left|\frac{a_{i+1}}{a_i}\right|\ \left(\overset{(3.26)}{\leq} M_i^{-2(m+1)^2+m+2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\leq M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}\right)$ in (3.33):

$$\begin{aligned}
a_ia_{i+2}\pm a_ia_{i+2}^{(1)}M_i^{-2(m+1)^2}\left|\frac{a_{i+1}}{a_i}\right|+a_i\frac{A_{i+2}^{(2)}(\xi)}{2!}M_i^{-4(m+1)^2}\left|\frac{a_{i+1}}{a_i}\right|^2\\
\leq\ M_i^{(m+1)^2}|A_{i+1}(\pm M_i^{-2(m+1)^2}\left|\frac{a_{i+1}}{a_i}\right|)|^2\leq 4M_i^{(m+1)^2}|a_{i+1}|^2,
\end{aligned}$$

by the previous result. It implies that

$$\begin{aligned}
\pm a_i a_{i+2}^{(1)} M_i^{-2(m+1)^2} \left|\frac{a_{i+1}}{a_i}\right| &\le 4M_i^{(m+1)^2}|a_{i+1}|^2 - a_i a_{i+2} \\
&\quad - a_i \frac{A_{i+2}^{(2)}(\xi)}{2!} M_i^{-4(m+1)^2} \left|\frac{a_{i+1}}{a_i}\right|^2 \\
&\le 4M_i^{(m+1)^2}|a_{i+1}|^2 + |a_i a_{i+2}| \\
&\quad + |a_i| \frac{|A_{i+2}^{(2)}(\xi)|}{2!} M_i^{-4(m+1)^2} \left|\frac{a_{i+1}}{a_i}\right|^2 \\
&\overset{(3.32)}{\le} 4M_i^{(m+1)^2}|a_{i+1}|^2 + |a_{i+1}|^2 \\
&\quad + M_i^{-3(m+1)^2+3}|a_{i+1}|^2 \\
&\le 6M_i^{(m+1)^2}|a_{i+1}|^2,
\end{aligned}$$

whence

$$|a_{i+2}^{(1)}| \le 6M_i^{3(m+2)^2}|a_{i+1}|.$$

Thus, we have for $|t| \le M_i^{-4(m+4)^2}\left|\frac{a_{i+1}}{a_i}\right| \left(\le M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|\right)$:

$$\begin{aligned}
|A_{i+2}(t)a_i| &= \left|a_i a_{i+2} + a_i a_{i+2}^{(1)} t + a_i \frac{A_{i+2}^{(2)}(\xi)}{2!} t^2\right| \\
&\le |a_i a_{i+2}| + |a_i||a_{i+2}^{(1)}||t| + |a_i| \frac{|A_{i+2}^{(2)}(\xi)|}{2!}|t|^2 \\
&\overset{(3.32)}{\le} |a_{i+1}|^2 + |a_i| 6M_i^{3(m+2)^2}|a_{i+1}| M_i^{-4(m+4)^2}\left|\frac{a_{i+1}}{a_i}\right| \\
&\quad + \frac{1}{2}|a_i| M_i^{(m+1)^2+4}|a_i| M_i^{-8(m+4)^2}\left|\frac{a_{i+1}}{a_i}\right|^2 \\
&\le 2|a_{i+1}|^2. \qquad (3.35)
\end{aligned}$$

For $|t| \le M_{i+1}^{-1}\left|\frac{a_{i+1}}{a_i}\right|$ and $j = 0, \dots, i-1$, consider

$$\begin{aligned}
|A_j(t)|\left|\frac{a_{i+1}}{a_i}\right|^{i+1-j} &= |A_j(t)|\left|\frac{a_{i+1}}{a_i}\right|\left|\frac{a_{i+1}}{a_i}\right|^{i-j} \\
&\overset{(3.26)}{\le} |A_j(t)|\left|\frac{a_{i+1}}{a_i}\right| M_i^{(m+2)(i-j)}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{i-j}{2}} \\
&\overset{(II.4_j)}{\le} M_i^{(m+2)(i-j)+1}|a_{i+1}| \\
&\le M_{i+1}|a_{i+1}|.
\end{aligned}$$

Now we are able to see that the index $i+1$ satisfies the conditions of I: $(I.1)$ and $(I.2)$ are clear. $(I.3)$ is true since $i \in II$. $(I.4_j)$, for $j = 0, \dots, i-1$, has been shown just above, $(I.4_i)$ and $(I.4_{i+1})$ are easy consequences of $(I.2)$ and $(I.3)$ (for $i+1$, respectively), and, finally, $(I.4_{i+2})$ corresponds to (3.35).

Consider now the subcase II'' where $|a_i a_{i+2}| \ge a_{i+1}^2$. Note that $a_{i+2} \ne 0$ (by lemma 3.4.4(4)).

Set $t = \pm M_i^{-2(m+1)^2}\left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} \left(\overset{(3.28)}{\le} M_i^{\frac{(m+1)^2+1}{2}-2(m+1)^2}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}} \le M_i^{-1}\left|\frac{a_i}{a_{i-1}}\right|^{\frac{1}{2}}\right)$

into (3.33), and we find

$$a_i a_{i+2} \pm a_i a_{i+2}^{(1)} M_i^{-2(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} + a_i \frac{A_{i+2}^{(2)}(\xi)}{2!} M_i^{-4(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right|$$
$$\leq \quad M_i^{(m+1)^2} \left|a_{i+1} \pm A_{i+1}^{(1)}(\eta) M_i^{-2(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}}\right|^2,$$

whence

$$\begin{aligned}
&\pm a_i a_{i+2}^{(1)} M_i^{-2(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} \\
\leq \quad & M_i^{(m+1)^2} a_{i+1}^2 \pm 2 M_i^{-(m+1)^2} a_{i+1} A_{i+1}^{(1)}(\eta) \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} \\
& + (A_{i+1}^{(1)}(\eta))^2 M_i^{-3(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right| - a_i a_{i+2} - a_i \frac{A_{i+2}^{(2)}(\xi)}{2!} M_i^{-4(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right| \\
\leq \quad & M_i^{(m+1)^2} a_{i+1}^2 + 2 M_i^{-(m+1)^2} |a_{i+1}| |A_{i+1}^{(1)}(\eta)| \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}} \\
& + (A_{i+1}^{(1)}(\eta))^2 M_i^{-3(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right| + |a_i a_{i+2}| \\
& + |a_i| \frac{|A_{i+2}^{(2)}(\xi)|}{2!} M_i^{-4(m+1)^2} \left|\frac{a_{i+2}}{a_i}\right| \\
\leq \quad & \Bigg(M_i^{(m+1)^2} + M_i^{-m^2-m+3} + \frac{1}{4} M_i^{-2m^2+2m+13} \\
& + 1 + \frac{1}{2} M_i^{-3m^2-6m+1} \Bigg) |a_i a_{i+2}| \\
\leq \quad & 5 M_i^{(m+1)^2},
\end{aligned}$$

using (3.31) and (3.32) in the one but last step. So we get

$$|a_{i+2}^{(1)}| \leq 5 M_i^{3(m+1)^2} |a_i a_{i+2}|^{\frac{1}{2}}.$$

For $|t| \leq M_{i+2}^{-1} \left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}}$, this gives

$$\begin{aligned}
|A_{i+2}(t) - a_{i+2}| \quad &= \quad \left| a_{i+2}^{(1)} t + \frac{A_{i+2}^{(2)}(\xi)}{2!} t^2 \right| \\
&\leq \quad |a_{i+2}^{(1)}| |t| + \frac{|A_{i+2}^{(2)}(\xi)|}{2!} |t|^2 \\
&\overset{(3.32)}{\leq} \quad 5 M_i^{3(m+1)^2} M_{i+2}^{-1} |a_{i+2}| + \frac{1}{2} M_i^{(m+1)^2+4} M_{i+2}^{-2} |a_{i+2}| \\
&\leq \quad \frac{1}{2} |a_{i+2}|,
\end{aligned}$$

whence $i+2$ satisfies $(II.2)$.

Finally, we investigate, whether $i+2$ fulfills $(II.4)$:

$$\begin{aligned}
|A_j(t)|\left|\frac{a_{i+2}}{a_i}\right|^{\frac{i+2-j}{2}} &= |A_j(t)|\left|\frac{a_{i+2}}{a_i}\right|\left|\frac{a_{i+2}}{a_i}\right|^{\frac{i-j}{2}}\\
&\overset{(3.28)}{\leq} |A_j(t)|\left|\frac{a_{i+2}}{a_i}\right|\left(M_i^{(m+1)^2+1}\left|\frac{a_i}{a_{i-2}}\right|\right)^{\frac{i-j}{2}}\\
&\overset{(II.4_j)}{\leq} M_i^{\frac{1}{2}((m+1)^2+1)(i-j)+1}|a_{i+2}|\\
&\leq M_{i+2}|a_{i+2}|
\end{aligned}$$

which works for $|t| \leq M_i^{-1}\left|\frac{a_i}{a_{i-2}}\right|^{\frac{1}{2}}$ and $j=0,\dots,i-1$. For $j=i$, the statement is trivial, and, for $j=i+1$, we observe that, if $|t| \leq M_{i+2}^{-1}\left|\frac{a_{i+2}}{a_i}\right|^{\frac{1}{2}}$:

$$\begin{aligned}
|A_{i+1}(t)| &= |a_{i+1}+A_{i+1}^{(1)}(\xi)t|\\
&\leq |a_{i+1}|+|A_{i+1}^{(1)}(\xi)||t|\\
&\overset{(3.31)}{\leq} |a_ia_{i+2}|^{\frac{1}{2}}+\frac{1}{2}M_i^{m+4}M_{i+2}^{-1}|a_ia_{i+2}|^{\frac{1}{2}}\\
&\leq 2|a_ia_{i+2}|^{\frac{1}{2}}.
\end{aligned}$$

Thus, we have just shown that the index $i+2$ belongs to II.

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 claim 2.

If $r-1$ satisfies the conditions of I, then inequality (3.17), with $j=1$, provides the estimate we are looking for. If $r-1$ satisfies the conditions of II, then inequality (3.29), with $j=1$, does it. Therefore, the lemma is proved. $\square$

The forgoing proof already implies the following variant of lemma 3.5.1.

LEMMA 3.5.2. *If in the preceding lemma 3.5.1 the inequality $A_{r-1}(0) \neq 0$ is replaced by the relations $A_{r-2}(0) \neq 0$ and $A_{r-1}(0) = 0$, then (3.15) has to be replaced by*

$$\left|\frac{A_r''(0)}{A_{r-2}(0)}\right| \leq M_r.$$

PROOF. Rereading the previous proof shows that now $r-2 \in J$. If $r-2$ belongs to I, then, by (3.22), we get

$$|A_r''(0)| \leq M_{r-2}^{2m+6}|A_{r-2}(0)| \leq M_r|A_{r-2}(0)|.$$

If $r-2$ belongs to I, then (3.32) gives

$$|A_r''(0)| \leq M_{r-2}^{(m+1)^2+4}|A_{r-2}(0)| \leq M_r|A_{r-2}(0)|.$$

Therefore, the proof is given. $\square$

Finally, we can formulate and prove the main result of this chapter.

THEOREM 3.5.3. *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 sequence-compact Hausdorff topological space, and the multiplicity of its roots does not exceed k. Furthermore, suppose that all $\frac{\partial^i}{\partial t^i}a_j(t,y)$ $(i=0,\dots,k; j=1,\dots,n)$ are continuous functions on

$(-1,1)\times\mathcal{M}$. Then, for any sequence-compact subset $K\subset(-1,1)\times\mathcal{M}$, there is a constant C_K such that, for all roots $x_j(t,y)$ ($j=1,\dots,n$) of P, there is the following estimate

$$\left|\frac{\partial}{\partial t}x_j(t,y)\right| < C_K \qquad \forall (t,y)\in K.$$

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 not bounded on a sequence-compact $K\subset(-1,1)\times\mathcal{M}$ for one $j\in\{1,\dots,n\}$. Without loss of generality let $j=1$, and we assume there is a sequence $(t_p,y_p)_{p\in\mathbb{N}}\subset K$ such that

$$\begin{aligned}
(t_p,y_p) &\overset{p\to\infty}{\longrightarrow} (t_\infty,y_\infty)\in K\\
x_1(t_p,y_p) &\overset{p\to\infty}{\longrightarrow} x_1(t_\infty,y_\infty)\\
\left|\frac{\partial}{\partial t}x_1(t_p,y_p)\right| &\overset{p\to\infty}{\longrightarrow} \infty.
\end{aligned}$$

Turn to a subsequence (again denoted by $(t_p,y_p)_p$) to obtain that the multiplicity of $x_1(t_p,y_p)$ equals, say, q for any $p\in\mathbb{N}$. Thus, $1\le q\le k$. Therefore, the multiplicity of $x_1(t_\infty,y_\infty)$, say s, satisfies $q\le s\le k$, for, if $s<q$, then the sequence $(x_1(t_p,y_p))_p\subset\mathbb{R}$ 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 $b_{p,1},\dots,b_{p,n}$ take the following form, by Taylor's formula,

$$b_{p,j}(t) = \frac{1}{(n-j)!}\left(\frac{\partial}{\partial x}\right)^{n-j}\Bigg|_{x=x_1(t_p,y_p)} P(t,y_p)(x).$$

Note that this is true for $b_{p,0}\equiv 1$, too.

Remember that, by theorem 3.2.1, $\frac{\partial}{\partial t}x_1(t_p,y_p)$ 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 ($b_{p,n-q}(t_p)\neq 0$ will be checked below). This would be contradictory to the assumption that $\frac{\partial}{\partial t}x_1(t_p,y_p)$ is unbounded (see lemma 3.4.3) and, therefore, finish the proof.

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 $|t - t_\infty| \leq 1$ and all $p \in \mathbb{N}$.
For next we observe that

$$b_{p,n-q}(t_p) = \frac{1}{q!} \left(\frac{\partial}{\partial x}\right)^q \Big|_{x=x_1(t_p,y_p)} P(t_p, y_p)(x) \neq 0$$

and, for $1 \leq j \leq q$,

$$b_{p,n-q+j}(t_p) = \frac{1}{(q-j)!} \left(\frac{\partial}{\partial x}\right)^{q-j} \Bigg|_{x=x_1(t_p,y_p)} P(t_p, y_p)(x) = 0,$$

the multiplicity of $x_1(t_p, y_p)$ being q.

Let us consider

$$\left(\frac{\partial}{\partial \tilde{x}}\right)^{q-j} Q_p(t)(\tilde{x}) \quad = \quad \sum_{i=0}^{n-q+j} \frac{(n-i)!}{(n-i-q+j)!} b_{p,i}(t) \tilde{x}^{n-i-q+j},$$

for $j = 1, 2$ and $j \leq q$. We want to apply lemma 3.5.1 or lemma 3.5.2 to this polynomial, j being 1 or 2, respectively. The correspondence between the present and the former used notation (up to unimportant constant factors) is the following:

$$\begin{array}{ccc} A_0 & \longleftrightarrow & b_{p,n-s} \\ A_1 & \longleftrightarrow & b_{p,n-s+1} \\ \vdots & & \vdots \\ A_{s-q+j} & \longleftrightarrow & b_{p,n-q+j}. \end{array}$$

Still to check are the differentiability conditions in lemma 3.5.1 and lemma 3.5.2. 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 clear, since $j \leq q$.
Thus, application of lemmas 3.5.1 and 3.5.2 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 $(t_p, y_p)_p \subset K$. Therefore,

$$\sup_{p \in \mathbb{N}} \left| \frac{b^{(j)}_{p,n-q+j}(t_p)}{b_{p,n-q}(t_p)} \right| < \infty \qquad (j = 1, 2). \tag{3.36}$$

In order to get (3.36) for $j = 3, \ldots, q$, too, consider

$$\left(\frac{\partial}{\partial x}\right)^{q-j} T_p(x) = \frac{q!}{j!} b_{p,n-q}(t_p) x^j + \cdots + \frac{(q-j)!}{j!} b^{(j)}_{p,n-q+j}(t_p).$$

To shorten notation we write $c_{p,j} = \frac{b^{(j)}_{p,n-q+j}(t_p)}{b_{p,n-q}(t_p)}$, for $j = 1, \ldots, q$. Applying lemma 3.4.4(7) to

$$\frac{j!}{q!} (b_{p,n-q}(t_p))^{-1} \left(\frac{\partial}{\partial x}\right)^{q-j} T_p(x) = x^j + \frac{j}{q} c_{p,1} x^{j-1} + \cdots + \frac{(q-j)!}{q!} c_{p,j},$$

gives

$$\begin{aligned} |c_{p,j}| \quad \leq \quad & C_2 \sum_{i=0}^{j-1} |c_{p,i}| \left(\min\left(\left|\frac{c_{p,j-1}}{c_{p,j-2}}\right|, \left|\frac{c_{p,j-1}}{c_{p,j-3}}\right|^{\frac{1}{2}} \right) \right. \\ & \left. + \min\left(\left|\frac{c_{p,j-2}}{c_{p,j-3}}\right|, \left|\frac{c_{p,j-2}}{c_{p,j-4}}\right|^{\frac{1}{2}} \right) \right)^{j-i}, \end{aligned}$$

where C_2 is a positive constant.
This formula allows to make induction on j. Suppose $c_{p,1}, \ldots, c_{p,j-1}$ are bounded

in p. If in one of the minima both expressions are unbounded, then both denominators have to converge to 0 (at least subsequences of it), since the nominators are bounded. But the denominators in both minima are consecutive coefficients in the hyperbolic polynomial $\frac{j!}{q!}(b_{p,n-q}(t_p))^{-1}\left(\frac{\partial}{\partial x}\right)^{q-j} T_p(x)$, whence, by lemma 3.4.4(4), $\lim_{p\to\infty} c_{p,j} = 0$ (the space of hyperbolic polynomials is closed). So, $c_{p,j}$ is bounded in p in any case. This finishes the induction.

Consequently, (3.36) is established for all $j = 1, \ldots, q$, and, hence, all coefficients of $(b_{p,n-q}(t_p))^{-1}T_p(x)$ are bounded in p. $\square$

CHAPTER 4

Wakabayashi's approach

In this chapter we shall present an approach to show the local boundedness of the derivatives of roots of a curve of hyperbolic polynomials which is due to Wakabayashi who published it in 1986, see [**41**]. It is shorter and more conceptual than Bronshtein's approach.

4.1. 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.

Lemma 4.1.1. *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 $j = 1, \ldots, n$. 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 $\mathrm{Im}(s) \leq 0$, $\left(1 + s\frac{d}{dx}\right) P(x) = 0$. Then, $1 + s\sum_{j=1}^n (x - \alpha_j)^{-1} = 0$. But this means that

$$\mathrm{Re}(s\sum_{j=1}^n (x - \alpha_j)^{-1}) = -1 \quad \text{and} \quad \mathrm{Im}(s\sum_{j=1}^n (x - \alpha_j)^{-1}) = 0. \tag{4.1}$$

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 (4.1) take the following form

$$-1 = \mathrm{Re}(su) = \mathrm{Re}(s)\mathrm{Re}(u) - \mathrm{Im}(s)\mathrm{Im}(u)$$

and

$$0 = \mathrm{Im}(su) = \mathrm{Re}(s)\mathrm{Im}(u) + \mathrm{Im}(s)\mathrm{Re}(u).$$

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 4.1.2. *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 $s \in \mathbb{R}$.*

The following lemma shows that, indeed, the operator $P \mapsto P + sP'$ reduces the multiplicity of the roots (see (4.2)), and it gives an estimate for the deviation the roots are subjected to under this operator (see (4.3)).

LEMMA 4.1.3. *Let $P(x) = \prod_{j=1}^{n}(x-\alpha_j^0)$ be a hyperbolic polynomial with roots $\alpha_1^0 \le \alpha_2^0 \le \cdots \le \alpha_n^0$. 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 $\alpha_1(s) \le \alpha_2(s) \le \cdots \le \alpha_n(s)$ and $\alpha_j(0) = \alpha_j^0$ ($j = 1, \dots, n$). Then, there exist positive constants $C_1(n)$ and $C_2(n)$, depending only on n, such that

$$\alpha_j(s) - \alpha_{j-1}(s) \ge C_1(n)|s| \quad \textit{for } s \in \mathbb{R} \textit{ and } 2 \le j \le n \tag{4.2}$$

and

$$0 < \pm(\alpha_j^0 - \alpha_j(s)) \le C_2(n)|s| \quad \textit{for } \pm s > 0 \textit{ and } 1 \le j \le n. \tag{4.3}$$

PROOF. First of all note that, for $s = 0$, (4.2) is trivial. Consider the case where $s > 0$. We make induction on the number of how often the operator $1 + s\frac{d}{dx}$ is applied to P: assume that, for a fixed $l \in \{1, \dots, n-1\}$, there is a positive constant $C_1(l)$ such that

$$\alpha_j^l(s) - \alpha_{j-1}^l(s) \ge C_1(l)s \quad \text{for } s > 0 \text{ and } 2 \le j \le l, \tag{4.4}$$

where $\left(1+s\frac{d}{dx}\right)^{l-1}P(t) = \prod_{j=1}^{n}(x-\alpha_j^l(s))$ and $\alpha_1^l(s) \le \alpha_2^l(s) \le \cdots \le \alpha_n^l(s)$. (4.4) is trivially satisfied, if $l = 1$. Put

$$\begin{aligned}
f(x,s) &= \frac{\left(1+s\frac{d}{dx}\right)^{l}P(x)}{\left(1+s\frac{d}{dx}\right)^{l-1}P(x)} = \frac{\left(1+s\frac{d}{dx}\right)\left(1+s\frac{d}{dx}\right)^{l-1}P(x)}{\left(1+s\frac{d}{dx}\right)^{l-1}P(x)} \\
&= 1 + s\frac{\frac{d}{dx}\left(1+s\frac{d}{dx}\right)^{l-1}P(x)}{\left(1+s\frac{d}{dx}\right)^{l-1}P(x)} \\
&= 1 + s\frac{\sum_{j=1}^{n}(x-\alpha_1^l(s))\cdots\widehat{(x-\alpha_j^l(s))}\cdots(x-\alpha_n^l(s))}{\prod_{j=1}^{n}(x-\alpha_j^l(s))} \\
&= 1 + s\sum_{j=1}^{n}(x-\alpha_j^l(s))^{-1}.
\end{aligned}$$

Since $s > 0$, we find, for $1 \le h \le n$ and $\alpha_{h-1}^l(s) < x < \alpha_h^l(s)$,

$$1 + sn(x-\alpha_1^l(s))^{-1} < f(x,s) < 1 + s(x-\alpha_1^l(s))^{-1} \quad \text{when } h = 1 \tag{4.5}$$

and

$$\begin{aligned}
&1 + s(x-\alpha_{h-1}^l(s))^{-1} + s(n-h+1)(x-\alpha_h^l(s))^{-1} < f(x,s) \\
&\qquad < A_h + s(x-\alpha_{h-1}^l(s))^{-1} + s(x-\alpha_h^l(s))^{-1} \quad \text{when } 2 \le h \le n,
\end{aligned} \tag{4.6}$$

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

$$\begin{aligned}
f(x,s) &= 1 + s\sum_{j=1}^{h-2} \underbrace{(x-\alpha_j^l(s))^{-1}}_{<(\alpha_{h-1}^l(s)-\alpha_{h-2}^l(s))^{-1}} + s(x-\alpha_{h-1}^l(s))^{-1} + s(x-\alpha_h^l(s))^{-1} \\
&\quad + s\sum_{j=h+1}^{n} \underbrace{(x-\alpha_j^l(s))^{-1}}_{<0} \\
&< A_h + s(x-\alpha_{h-1}^l(s))^{-1} + s(x-\alpha_h^l(s))^{-1}.
\end{aligned}$$

We assert that, for $1 \le h \le n$ and $\alpha_{h-1}^l(s) < \alpha_h^l(s)$, this yields

$$\begin{cases}
\alpha_{h-1}^l(s) < \alpha_h^{l+1}(s) < \alpha_h^l(s), \\
\alpha_1^l(s) - sn < \alpha_1^{l+1}(s) < \alpha_1^l(s) - s \quad \text{when } h = 1, \\
\alpha_h^l(s) - \frac{1}{2}\Big(X_h + s(n-h+2) - [(X_h - s(n-h+2))^2 + 4sX_h]^{\frac{1}{2}}\Big) \\
\quad < \alpha_h^{l+1}(s) < \alpha_h^l(s) - \frac{1}{2}F(X_h, \frac{2s}{A_h}) \quad \text{when } 2 \le h \le n,
\end{cases} \tag{4.7}$$

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 (4.7) 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 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 (4.7), since the second inequality is analogous.

The second row of (4.7) is obtained by applying (4.5) 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 (4.6), by putting $x = \alpha_h^{l+1}(s)$. Thus, the assertion is shown.

Since $(X_h-(n-h+2)s)^2+4sX_h = (X_h-(n-h)s)^2+4(n-h+1)s^2 \ge (X_h-(n-h)s)^2$, (4.7) yields

$$0 < \alpha_h^l(s) - \alpha_h^{l+1}(s) < (n-h+1)s, \text{ for } 1 \le h \le n. \tag{4.8}$$

Moreover, (4.7) 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 (4.4), $X_h=\alpha_h^l(s)-\alpha_{h-1}^l(s)\geq C_1(l)s$, and $A_h=1+s(h-2)(\alpha_{h-1}^l(s)-\alpha_{h-2}^l(s))^{-1}\leq 1+(h-2)C_1(l)^{-1}$. This and the properties of F to be positively homogeneous and to satisfy $F(u_1,v_1)\geq F(u_2,v_2)$, 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 (4.4) is valid, replacing 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 (4.2), for $s>0$.
To get (4.3), for $s>0$, note that, by definition, $\alpha_j^0=\alpha_j^1(s)$ and $\alpha_j^n(s)=\alpha_j(s)$. Then, (4.8) 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 $j=1,\dots,n$. This completes the proof, when $s>0$. Similarly, one can prove the lemma when $s<0$. $\square$

Furthermore, we shall need the following lemma.

LEMMA 4.1.4. *Consider $P(x)=x^n+\sum_{j=1}^n a_jx^{n-j}$ and $Q(x)=\sum_{j=1}^n b_jx^{n-j}$, and write $P(x)+Q(x)=\prod_{j=1}^n(x-\alpha_j(b_1,\dots,b_n))$, where $\alpha_1,\dots,\alpha_n$ are continuous functions of $(b_1,\dots,b_n)\in\mathbb{C}^n$. Then there exists a positive constant $C(n)$, depending only on n, such that*

$$|\alpha_j(b_1,\dots,b_n)-\alpha_j^0|\leq C(n)\max_{1\leq k\leq n}\left(|b_k|^{\frac{1}{k}}+|b_k|^{\frac{1}{n}}|\alpha_j^0|^{1-\frac{k}{n}}\right),\quad 1\leq j\leq n,\tag{4.9}$$

where $\alpha_j^0=\alpha_j(0,\dots,0)$, $1\leq j\leq n$, are the roots of P.

PROOF. We shall prove that 4.9 holds for $j=1$. The remaining cases $j=2,\dots,n$ are identical. There is a integer k_0 with $1\leq k_0\leq n$ such that none of $\alpha_2^0,\dots,\alpha_n^0$ lies in $U:=\{z\in\mathbb{C}:(k_0-1)A\leq|z-\alpha_1^0|<k_0A\}$, 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 1.2.1. 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 $1 \leq i \leq n$, $A \geq C'(n)\left(|b_i|^{\frac{1}{i}} + |b_i|^{\frac{1}{n}}|\alpha_1^0|^{1-\frac{i}{n}}\right)$ 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

$$|P(z)| - |Q(z)| \geq n|b_i|\left(|\alpha_1^0| + (k_0 - 2^{-1})A\right)^{n-i} - \sum_{j=1}^{n} |b_j||z|^{n-j},$$

where we choose i such that $|b_i|$ is maximal. Since $|\alpha_1^0| + (k_0 - 2^{-1})A \geq |z|$, it implies that on the circle $|z - \alpha_1^0| = (k_0 - 2^{-1})A$ the inequality $|P(z)| > |Q(z)|$ holds. Applying Rouché's theorem 1.2.1, we obtain

$$|\alpha_1(b_1, \ldots, b_n) - \alpha_1^0| \leq (k_0 - 2^{-1})A \leq (n - 2^{-1})A,$$

and, if we put $A := C'(n)\max_{1\leq k\leq n}\left(|b_k|^{\frac{1}{k}} + |b_k|^{\frac{1}{n}}|\alpha_1^0|^{1-\frac{k}{n}}\right)$, then the lemma is proved for $j = 1$. □

We will make use of the following lemma, too.

Lemma 4.1.5. *Let M be an arc-wise connected subset of $\mathbb{R}^n$, U a Hausdorff topological space, and $S = \{s \in \mathbb{C} : |s| \leq s_0$ and $\mathrm{Im}(s) \leq 0\}$, 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) tells us that $f(\_, w_1, u_1)$ cannot vanish identically on S. So s_1 is an isolated root of $f(\_, w_1, u_1)$, and we may find a small circle C centered at s_1 such that s_1 is the only zero of $f(\_, w_1, u_1)$ inside and on C. 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 $s \in C$. Application of Rouché's theorem 1.2.1 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 conditions (i) - (iii), we can apply theorem 1.2.2 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 $\mathrm{int}S$, 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. □

4.2. Microhyperbolicity

In this section we shall make a short excursion to the theory of partial differential equations, where the notion of microhyperbolicity appears and plays a key part in questions related to the Cauchy problem. Note that the study of hyperbolic polynomials is mostly motivated by this background, and, in fact, Bronshtein and Wakabayashi are settled in this mathematical area. Nevertheless, we will not enter deeply the related theory but only introduce the notion of microhyperbolicity and discuss a few properties. The following considerations are based on [**14**] and [**15**].

Definition 4.2.1. 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

$$F(x + it\Theta) \neq 0, \quad \text{if } 0 < t \leq 1 \text{ and } x \in U. \tag{4.10}$$

Lemma 4.2.2. *If F satisfies* (4.10) *and $F(x_0 + t\Theta)$ has a zero of multiplicity m exactly when $t = 0$, where $x_0 \in U$, then*

$$F(x) = F_0(x) + O(|x - x_0|^{m+1}), \quad \text{for } x \in U,$$

where F_0 is a homogeneous polynomial of degree m with

$$F_0(\Theta) \neq 0 \quad \text{and} \quad F_0(x + it\Theta) \neq 0, \text{ if } t \in \mathbb{R}\backslash\{0\} \text{ and } x \in \mathbb{R}^n. \tag{4.11}$$

Proof. To simplify notation we assume without loss of generality that $0 \in U$ and $x_0 = 0$. Note that, if $m = 0$, i.e., $F(t\Theta)$ does not vanish at $t = 0$, then the lemma is trivial. Thus, suppose $m \geq 1$. 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 $\lambda \geq \frac{1}{m}$, since $g(0, s) = F(sy)$ vanishes at $s = 0$, and $\lambda < 1$, for if we had $\lambda \geq 1$ then in particular $g(t, s) = O(|t| + |s|)^m$ would follow. 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 $at^j s^k$ is a term in the expansion of $g(t, s)$ with $j + \frac{k}{\lambda} = m$, then q divides $m - j$, because $m - j = \frac{k}{\lambda} = \frac{kq}{p}$ implies $p(m - j) = kq$ and, since p and q are relatively prime, the statement follows. 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 nonzero 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 $1 \le p < q$. On the other hand the roots of $g_0^{\pm}(w) = 0$ satisfy $\mathrm{Im}(w) \le 0$, for if $g(w|s|^\lambda, s) = F(\mathrm{Re}(w)|s|^\lambda\Theta + sy + i\mathrm{Im}(w)|s|^\lambda\Theta) = 0$ and s is sufficiently small, then, by (4.10), $\mathrm{Im}(w)$ cannot be positive. So the assumption $\lambda \ne 1$ 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 (4.11), namely $F_0(\Theta) \ne 0$, 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 (4.10), yields $F_0(x + w\Theta) \ne 0$, if $x \in \mathbb{R}^n$ and $\mathrm{Im}(w) > 0$. Hence $F_0(x + w\Theta) = (-1)^m F_0(-x - w\Theta) \ne 0$, if $x \in \mathbb{R}^n$ and $\mathrm{Im}(w) < 0$. This shows the second part of (4.11) and completes the proof. $\square$

The polynomial F_0 appearing in the previous lemma is often referred to as the *localization polynomial*.

4.3. Lipschitz continuity of the roots

The following theorem provides a variant of Bronshtein's theorem 3.5.3.

THEOREM 4.3.1. *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 $\partial_t^\alpha a_j(t,y)$ ($|\alpha| \le k, 1 \le j \le n$) are continuous on $T \times \mathcal{Y}$ and that there exist constants $C > 0$ and $0 < \delta \le 1$ such that*

$$|\partial_t^\alpha a_j(t,y) - \partial_t^\alpha a_j(t',y)| \le C|t - t'|^\delta, \tag{4.12}$$

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 $\lambda_1(t,y) \le \dots \le \lambda_n(t,y)$ of $P(t,y)$ satisfy

$$|\lambda_j(t,y) - \lambda_j(t',y)| \le C_U|t - t'|^r, \tag{4.13}$$

for $1 \le j \le n$, $t, t' \in U$ and $y \in \mathcal{Y}$, where $r = \min\left\{1, \frac{k+\delta}{n}\right\}$.

PROOF. 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 $\mathrm{Im}(z) \le 0$, where we demand $z^r = |z|^r e^{-ir\pi}$ if $z \le 0$. Corollary 4.1.2 tells us 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 4.1.3, if $z \ge 0$, or $z \in \mathbb{R}$ and $r = 1$, then we find positive constants $C_1(n)$ and $C_2(n)$ such that

$$\tilde{\lambda}_j(t,y,z) - \tilde{\lambda}_{j-1}(t,y,z) \ge C_1(n)|z|^r, \quad 2 \le j \le n \tag{4.14}$$

and

$$|\lambda_j(t,y) - \tilde{\lambda}_j(t,y,z)| \le C_2(n)|z|^r, \quad 1 \le j \le n, \tag{4.15}$$

for $(t,y) \in T \times \mathcal{Y}$, where $\tilde{\lambda}_1(t,y,z) \le \tilde{\lambda}_2(t,y,z) \le \dots \le \tilde{\lambda}_n(t,y,z)$ are the ordered roots of $\tilde{P}(t,y,z)$.

If $z \leq 0$ and $r < 1$, then $\operatorname{Im}(z^r) = -|z|^r \sin r\pi \leq 0$. Therefore, $P(t,y)(x+z^r) \neq 0$, if $0 > \operatorname{Im}(x+z^r) = \operatorname{Im}(x) + \operatorname{Im}(z^r) = \operatorname{Im}(x) - |z|^r \sin r\pi$, since $P(t,y)$ is hyperbolic. Application of lemma 4.1.1 shows that

$$\tilde{P}(t,y,z)(x+z^r) \neq 0, \quad \text{when } \operatorname{Im}(x) < |z|^r \sin r\pi. \tag{4.16}$$

Next, for $1 \leq j \leq n$, let us expand in means of Taylor's formula

$$a_j(t+z\xi,y) = \sum_{|\alpha|\leq k} \frac{\partial_t^\alpha a_j(t,y)}{\alpha!} z^{|\alpha|}\xi^\alpha + \tilde{a}_j(t,\xi,y,z),$$

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 $0 \leq \vartheta \leq 1$,

$$\begin{aligned}
|\tilde{a}_j(t,\xi,y,z)| &= \left| \sum_{|\alpha|=k} \frac{\partial_t^\alpha a_j(t+\vartheta z\xi,y)}{\alpha!} z^{|\alpha|}\xi^\alpha - \sum_{|\alpha|=k} \frac{\partial_t^\alpha a_j(t,y)}{\alpha!} z^{|\alpha|}\xi^\alpha \right| \\
&\leq \sum_{|\alpha|=k} |\partial_t^\alpha a_j(t+\vartheta z\xi,y) - \partial_t^\alpha a_j(t,y)| \frac{|z|^{|\alpha|}|\xi_1|^{\alpha_1}\cdots|\xi_m|^{\alpha_m}}{\alpha!} \\
&\overset{(4.12)}{\leq} A|z|^{k+\delta}|\xi|^{k+\delta},
\end{aligned}$$

for a positive constant A, by the assumptions of the theorem and since $|\xi_1|^{\alpha_1}\cdots|\xi_m|^{\alpha_m} \leq (\max_{1\leq j\leq m}|\xi_j|)^{|\alpha|} \leq K|\xi|^{|\alpha|}$ (equivalence of norms). If, moreover, $|z| \leq 1$ and $|\xi| \leq 1$, this gives

$$|\tilde{a}_j(t,\xi,y,z)| \leq A|z|^{nr}|\xi|^{nr}, \tag{4.17}$$

remembering that $nr = \min\{n, k+\delta\}$.

Let U be an open relative-compact subset of T, and define

$$Q(t,\xi,y,z)(x) = \left(1 + z^r \frac{\partial}{\partial x}\right)^{n-1} \left(x^n + \sum_{j=1}^n \left(\sum_{|\alpha|\leq k} \frac{\partial_t^\alpha a_j(t,y)}{\alpha!} z^{|\alpha|}\xi^\alpha\right) x^{n-j}\right).$$

Now, (4.14) tells us 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 (4.14) does not appear in the right-hand side of (4.14). Observe that

$$\begin{aligned}
Q(t,\xi,y,z)(x) &= \left(1 + z^r \frac{\partial}{\partial x}\right)^{n-1} \cdot \\
&\quad \cdot \left(x^n + \sum_{j=1}^n \left(a_j(t+z\xi,y) - \tilde{a}_j(t,\xi,y,z)\right) x^{n-j}\right) \\
&= \tilde{P}(t+z\xi,y,z)(x) - \left(1 + z^r \frac{\partial}{\partial x}\right)^{n-1} \sum_{j=1}^n \tilde{a}_j(t,\xi,y,z) x^{n-j}.
\end{aligned}$$

Now, if we choose z and ξ sufficiently small, then, by (4.17), 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 4.1.4. Otherwise put, there are positive constants δ_0 and δ_1 such that $Q(t,\xi,y,z)(x) = 0$ has only simple roots, for $(t,\xi,y) \in \Omega(U;\delta_1) := \{(t,\xi,y) \in U \times \mathbb{R}^m \times \mathcal{Y} : |\xi| \leq \delta_1\}$, if $0 < z \leq \delta_0$ or $0 < |z| \leq \delta_0$ 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 2.2.1. 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 $0<z\le\delta_0$ or $0<|z|\le\delta_0$ 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 2.2). Note that $Q(t,\xi,y,0)=P(t,y)$ whose roots are all real by assumption. Summarizing we find that $Q(t,\xi,y,z)(x)=0$ has only real roots, for $(t,\xi,y)\in\Omega(U;\delta_1)$, if $0\le z\le\delta_0$ or $-\delta_0\le z\le\delta_0$ 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 $0\le z\le\delta_0$, where we assume that $\Lambda_1(t,\xi,y,z)\le\Lambda_2(t,\xi,y,z)\le\cdots\le\Lambda_n(t,\xi,y,z)$. Consider the second term on the right-hand side of

$$\begin{aligned}Q(t,\xi,y,z)(x+z^r) \quad &= \quad \tilde{P}(t+z\xi,y,z)(x+z^r)\\ &\qquad -\left(1+z^r\frac{\partial}{\partial x}\right)^{n-1}\sum_{j=1}^{n}\tilde{a}_j(t,\xi,y,z)(x+z^r)^{n-j},\end{aligned}\tag{4.18}$$

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 4.1.4, 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 (4.17), we find

$$|b_{n-i}| \quad\le\quad \sum_{j=1}^{n-i}\sum_{k=i}^{n-j}\frac{k!}{i!}\binom{n-j}{k}\binom{n-1}{k-i}z^{(n-i-j)r}Az^{nr}|\xi|^{nr}.$$

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 $0\le z\le\delta_0$. Therefore, the application of lemma 4.1.4 gives

$$|\Lambda_j(t,\xi,y,z)-(\tilde{\lambda}_j(t+z\xi,y,z)-z^r)|\le kz^r,$$

with a constant k, since $\tilde{\lambda}_j(t+z\xi,y,z)-z^r$ (taking the part of α_j^0 in lemma 4.1.4) 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

$$|\Lambda_j(t,\xi,y,z)-\tilde{\lambda}_j(t+z\xi,y,z)|\le cz^r,\quad 1\le j\le n,\tag{4.19}$$

for $(t,\xi,y)\in\Omega(U;\delta_1)$ and $0\le z\le\delta_0$.
Moreover, we claim

$$Q(t,\xi,y,z)(x+z^r)\neq 0,\tag{4.20}$$

for $(t,\xi,y)\in\Omega(U;\delta_1)$, $\mathrm{Im}(x)<0$ and $-\delta_0\le z\le\delta_0$, modifying δ_0 and δ_1, if necessary. In fact, if $r=1$ or $0\le z\le\delta_0$, then the equation $Q(t,\xi,y,z)(x+z^r)=0$ has only real roots, as we found before. Still to investigate is the case, when $r<1$ and $-\delta_0\le z<0$. Then, we see, by (4.16), 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 (4.18), by lemma 4.1.4 and by (4.17), 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.

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:

$$Q(t, \xi, y, z + s\zeta)(x + z^{r-1}s\tau + (z + s\zeta)^r) = s^\mu \left(Q_{(x,z;t,\xi,y)}(\tau, \zeta) + o(1)\right), \quad \text{as } s \searrow 0,$$

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

$$Q_{(x,z;t,\xi,y)}(-1, 0) \neq 0 \quad \text{and} \quad Q_{(x,z;t,\xi,y)}(\tau, \zeta) \neq 0, \text{ if } \mathrm{Im}(\tau) < 0 \text{ and } \zeta \in \mathbb{R}. \tag{4.21}$$

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 (4.20), we find $Q(t, \xi, y, z + s\zeta)(x + z^{r-1}s\tau + (z + s\zeta)^r) = Q(t, \xi, y, z + s\zeta)(x + z^{r-1}s\mathrm{Re}(\tau) + iz^{r-1}s\mathrm{Im}(\tau) + (z + s\zeta)^r) \neq 0$, since $\mathrm{Im}(x + z^{r-1}s\tau) = z^{r-1}s\mathrm{Im}(\tau) < 0$. (The part of the parameter t in (4.10) is played here by $-z^{r-1}s\mathrm{Im}(\tau)$). Lemma 4.2.2 yields the existence of the localization and (4.21). Note that $Q_{(x,z;t,\xi,y)}(\tau, \zeta)$ can be defined and fulfills (4.21), if $r = 1$ and $z = 0$, too.

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 $r = 1$:

Let $r < 1$. We are going to check now, whether this f satisfies the assumptions of lemma 4.1.5. f is clearly holomorphic in s, for $\mathrm{Im}(s) < 0$ (it corresponds to (i) in lemma 4.1.5). We find $f(s, \zeta, (x, t, \tau, \xi, y, z)) \neq 0$, if $\mathrm{Im}(x) < 0$ and $s \in \mathbb{R}$, since then $z + s\zeta \in [-\delta_0, \delta_0]$ and $\mathrm{Im}(x + z^{r-1}s\tau) = \mathrm{Im}(x) < 0$, and, by (4.20), the statement follows. 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 $z \in (0, \epsilon]$. 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 $\mathrm{Im}(s) < 0$. It follows from (4.20), 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 4.1.5 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

$$Q(t, \xi, y, z + s\zeta)(x + z^{r-1}s\tau + (z + s\zeta)^r) \neq 0, \tag{4.22}$$

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 $z \in (0, \epsilon]$.

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: $Q_{(x_0,0;t_0,\xi_0,y_0)}(\tau_0, 0) = (-\tau_0)^\mu Q_{(x_0,0;t_0,\xi_0,y_0)}(-1, 0) \neq 0$, by (4.21), and continuity in the second variable gives the statement. The constant c' 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 $\tau_0^{-1}\zeta < \zeta \leq c'(1)$. 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

$$Q(t, \xi, y, z + s\xi)(x + s\tau + (z + s\zeta)) \neq 0, \tag{4.23}$$

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:

$$Q(t, \xi, y, z + s\xi)(x + s\tau + (z + s\zeta)) = \sum_{j=0}^{\mu_0} Q_j(t, \xi, y, z, \tau, \zeta)(x) s^j + o(s^{\mu_0}),$$

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 $j < \mu_0$. Assertion (4.23) follows by continuity.

Let us now apply lemma 4.1.5 again: (4.23) corresponds to condition (iii); (i), (ii) and (iv) are obvious, since (4.20) holds for $r = 1$, too. Therefore,

$$Q(t, \xi, y, z + s\xi)(x + z^{r-1} s\tau + (z + s\zeta)^r) \neq 0, \tag{4.24}$$

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 $z \in [0, \epsilon]$.

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 (4.24) holds. Since $\bar{U}$ 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

$$Q(t, \xi, y, z + s\xi)(x + z^{r-1} s\tau + (z + s\zeta)^r) \neq 0, \tag{4.25}$$

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 $\tau_0 \geq 1$,

$$Q_{(x, z; t, \xi, y)}(\tau_0, \zeta) \neq 0, \tag{4.26}$$

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 (4.25), 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 1.2.1 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.

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

$$\begin{aligned} 0 &= Q(t, \xi, y, z + s\zeta)(\Lambda_j(t, \xi, y, z + s\zeta) + (z + s\zeta)^r) \\ &= s^{\mu} \big(Q_{(\Lambda_j(t, \xi, y, z), z; t, \xi, y)}(z^{1-r} s^{-1}(\Lambda_j(t, \xi, y, z + s\zeta) - \Lambda_j(t, \xi, y, z)), \zeta) \\ &\quad + o(1) \big), \end{aligned} \tag{4.27}$$

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 (4.27) 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 (4.26) this is impossible, if $z^{1-r} \frac{\partial}{\partial s}\big|_{s=0} \Lambda_j(t, \xi, y, z + s\zeta) \geq 1$. So we have found

$$\frac{\partial}{\partial s}\bigg|_{s=0} \Lambda_j(t, \xi, y, z + s\zeta) < z^{r-1}, \tag{4.28}$$

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 $\Lambda_j(t,\xi,y,0) = \lambda_j(t,y)$)

$$\begin{aligned}
\lambda_j(t+z\xi,y) - \lambda_j(t,y) \quad &= \quad \lambda_j(t+z\xi,y) - \tilde{\lambda}_j(t+z\xi,y,z) \\
&\quad\quad + \tilde{\lambda}_j(t+z\xi,y,z) - \tilde{\lambda}_j(t,y,z) \\
&\quad\quad + \tilde{\lambda}_j(t,y,z) - \lambda_j(t,y) \\
&\overset{(4.15)}{\leq} \quad 2C_2(n)z^r + \tilde{\lambda}_j(t+z\xi,y,z) - \tilde{\lambda}_j(t,y,z) \\
&= \quad 2C_2(n)z^r + \tilde{\lambda}_j(t+z\xi,y,z) - \Lambda_j(t,\xi,y,z) \\
&\quad\quad + \Lambda_j(t,\xi,y,z) - \Lambda_j(t,\xi,y,0) \\
&\quad\quad + \Lambda_j(t,\xi,y,0) - \tilde{\lambda}_j(t,y,z) \\
&\overset{(4.15),(4.19)}{\leq} \quad 3C_2(n)z^r + cz^r \\
&\quad\quad + \Lambda_j(t,\xi,y,z) - \Lambda_j(t,\xi,y,0) \\
&\leq \quad 3C_2(n)z^r + cz^r \\
&\quad\quad + kz\, \frac{\partial}{\partial s}\bigg|_{s=0} \Lambda_j(t,\xi,y,z+s\zeta) \quad \text{(for some } \zeta) \\
&\overset{(4.28)}{\leq} \quad K'z^r,
\end{aligned}$$

for $(t,\xi,y) \in \Omega(U;\delta_1)$ and $z \in [0,\epsilon]$, where k and K' are positive constants. Exchanging $t+z\xi$ and t in the previous calculation, and recalling the compactness of $\bar{U}$ 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 (4.13) and completes the proof of the theorem. □

Note that theorem 4.3.1 implies Bronshtein's result in theorem 3.5.3: 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, 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 $\frac{\partial}{\partial t}\lambda_j(t,y)$ is bounded on U.

CHAPTER 5

An application of Bronshtein's result

5.1. 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 polynomials with coefficients in C^{3n}. This conclusion is best possible, since the second derivatives may be unbounded even if the coefficients are smooth: e.g. consider $P(t,y)(x) = x^2 - t^2 - y^2 = 0$ with $y \in [-1,1]$, then the second derivatives of the roots $\frac{\partial^2}{\partial t^2}x(t,y) = \pm y^2(t^2+y^2)^{-\frac{3}{2}}$ are unbounded. The following theorem is due to Kriegl, Losik and Michor [**17**].

THEOREM 5.1.1. *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 (proposition 2.4.1). By replacing x by $y = x - \frac{a_1(t)}{n}$, we may assume that $a_1 \equiv 0$.
According to the strong multiplicity lemma 2.3.8, for $m \ge n$ 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 2.5.2.
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

$$P^1(t)(z) = z^n + a_{2,2}(t)z^{n-2} - a_{3,3}(t)z^{n-3} + \cdots + (-1)^n a_{n,n}(t)$$

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 3.5.3 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 $a_2(0) \ne 0$, then we apply the splitting lemma 2.3.3: 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 equal to, say, c_i, and where the c_i are distinct. By the argument above applied to each P_i seperately, 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 $i\in\{1,\dots,n\}$ be fixed. For $t_m\to 0$ there are $k_m\in\{1,\dots,n\}$ 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 $j_m\in\{1,\dots,n\}$ 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 2.5.2, is indeed C^1, and (1) is proved.

To the proof of (2): Let all a_i be C^{3n}. Remember that $a_1\equiv 0$.

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 $P^1(t)$ 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 2.3.3 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 degree 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), 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 $t \geq t_0$ 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 $x_1, \ldots, x_n$ 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), 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 $(t_m)_{m \in \mathbb{R}}$ 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 $\{1, \ldots, n\}$ not depending on m. Consequently, $y'_i(t)$ has at most $x'_1(t_0), \ldots, x'_n(t_0)$ 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) 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) as just used. 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 $x = (x_1, \ldots, x_n)$ 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 $y = (y_1, \ldots, y_n)$ 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 2.1.1, where we treated the quadratic case, shows that the differentiability assumptions for the curve of polynomials P in theorem 5.1.1 can possibly be improved.

Part 2

Lifting smooth curves over invariants

CHAPTER 6

Isometric action of Lie groups and invariants

6.1. A different point of view

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 $\mathbb{R}$ smoothly, real analytically or continuously differentiably of a certain degree. And we investigated the problem of finding parameterizations by t of the roots of $P(t)$ with best possible differentiability properties. Note that in section 2.7 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.

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. 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, namely $\sigma(\mathbb{R}^n)$, and the lift $\bar{c}$ corresponds to a parameterization of the roots of P.

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, two heads, "\sigma"] & \sigma(\mathbb{R}^n) \subseteq \mathbb{R}^n \\
& \mathbb{R} \arrow[u, "c"'] \arrow[ul, dotted, "\exists \bar{c}\,?"]
\end{tikzcd}$$

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 the first remark after theorem 6.2.3 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 $\sigma \circ \bar{c} = c$?

$$\begin{tikzcd} V \arrow[r, two heads, "\sigma"] & \sigma(V) = V/G \subseteq \mathbb{R}^n \\ & \mathbb{R} \arrow[u, "c"'] \arrow[ul, dotted, "\exists \bar{c}\,?"] \end{tikzcd}$$

Reformulation and generalization as presented here are taken from [**2**].

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.

6.2. The space $\sigma(V)$

Remember the characterization of the space of hyperbolic polynomials with a fixed degree, given in theorem 2.2.1. There is a similar description of the orbit space of an arbitrary orthogonal representation of a compact Lie group. Before we dedicate our attention to this description let us concentrate on the setting. It will be fixed throughout the remaining chapters.

Let G be a compact Lie group and let $\rho : G \to O(V)$ be an orthogonal representation in a real finite dimensional Euclidean vector space V with inner product $\langle .|. \rangle$. By a classical theorem of Hilbert and Nagata the algebra $\mathbb{R}[V]^G$ of invariant polynomials on V is finitely generated, see e.g. [**27**], [**42**] for details. 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 $t \in \mathbb{R}$. Therefore, the pre-image of 0 under σ consists only of 0: $\sigma^{-1}(0) = \{0\}$. 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.

LEMMA 6.2.1. *In the above situation we have:*

(1) *σ is proper, i.e., pre-images of compact sets are compact.*
(2) *σ seperates orbits of G.*
(3) *There is a map $\sigma' : V/G \to \mathbb{R}^n$ such that the following diagram commutes,*

$$\begin{tikzcd} V \arrow[r, "\sigma"] \arrow[d, "\pi"'] & \mathbb{R}^n \\ V/G \arrow[ur, dotted, "\sigma'"'] & \end{tikzcd}$$

and σ' is a homeomorphism onto its image.

PROOF. To (1): Consider $\sigma_1 \in \mathbb{R}[V]^G$ from above. By the theorem of Hilbert and Nagata, see e.g. [**27**] or [**42**], there is a polynomial $p \in \mathbb{R}[\mathbb{R}^n]$ such that $\sigma_1(x) = p(\sigma(x))$. If $(x_m)_m \subseteq V$ is an unbounded sequence, then $(\sigma_1(x_m))_m$ 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: $g.x = h.y$ implies $G.x = G.y$. 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: let O be open in $\mathbb{R}^n$, then $\sigma^{-1}(O)$ is open in V and we have $\pi^{-1}(\sigma'^{-1}(O)) = \pi^{-1}(\{\pi(v) : \sigma(v) \in O\}) = \sigma^{-1}(O)$. So on every compact subset of V/G we know that σ' is a homeomorphism onto its image, since the involved spaces are Hausdorff. Now take an arbitrary diverging sequence in V/G. It is the image under π of some equally diverging sequence in V. If this sequence in V has an unbounded subsequence, then by (1), its image under σ is unbounded as well, in particular divergent. If instead the diverging sequence in V (therefore its image under π, our starting sequence) is bounded, then it is contained in a compact subset of V, our starting sequence is contained in a compact subset of V/G, and here σ' is a homeomorphism, as we have noted. Consequently, its image under σ' is divergent as well. So we have shown that a sequence in V/G is convergent iff its image under σ' in $\mathbb{R}^n$ is convergent and, with that, that σ' is a homeomorphism onto its image. □

In the sequel we shall identify V/G and $\sigma(V)$ via the homeomorphism σ' given in lemma 6.2.1.

Note that if a Lie group G is acting smoothly on a manifold 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 iff $f \circ \pi : M \to \mathbb{R}$ 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:

THEOREM 6.2.2. [**27**], [**39**]
Consider a finite dimensional representation $\rho : G \to O(V)$ 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 surjective.

Let $\langle .|. \rangle$ denote also the G-invariant dual inner product on the dual space V^*. The differentials $d\sigma_i : V \to V^*$ are G-equivariant:

$$\begin{aligned} \langle z, g.d\sigma_i(x) \rangle &= \langle g.z, d\sigma_i(x) \rangle = d\sigma_i(g.x).dl_g(x).z \\ &= d(\sigma_i \circ l_g)(x).z = d\sigma_i(g.x).z = \langle z, d\sigma_i(g.x) \rangle, \end{aligned}$$

for arbitrary z, where $l_g : V \to V$ denotes the left-action by the element $g \in G$. Therefore, the polynomials $v \mapsto \langle d\sigma_i(v) | d\sigma_j(v) \rangle$ are in $\mathbb{R}[V]^G$, and they are entries of an $n \times n$ symmetric matrix valued polynomial

$$B(v) := \begin{pmatrix} \langle d\sigma_1(v) | d\sigma_1(v) \rangle & \ldots & \langle d\sigma_1(v) | d\sigma_n(v) \rangle \\ \vdots & \ddots & \vdots \\ \langle d\sigma_n(v) | d\sigma_1(v) \rangle & \ldots & \langle d\sigma_n(v) | d\sigma_n(v) \rangle \end{pmatrix}.$$

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 2.2: Then $G = S_n$ acts on $V = \mathbb{R}^n$ by permuting the coordinates, and $\mathbb{R}[V]^G = \mathbb{R}[\sigma_1, \ldots, \sigma_n]$, where $\sigma_1, \ldots, \sigma_n$ are the elementary symmetric polynomials which are algebraically independent, whence $V /\!\!/ G = \mathbb{R}^n$. We may choose different generators $s_1, \frac{1}{2}s_2, \ldots, \frac{1}{n}s_n$ of $\mathbb{R}[V]^G$, where the $s_i(x_1, \ldots, x_n) = \sum_{j=1}^n x_j^i$ are the Newton polynomials. Then, for $x = (x_1, \ldots, x_n)$,

$$\begin{aligned} \langle d(\frac{1}{i}s_i)(x) | d(\frac{1}{j}s_j)(x) \rangle &= \langle (x_1^{i-1}, x_2^{i-1}, \ldots, x_n^{i-1}) | (x_1^{j-1}, x_2^{j-1}, \ldots, x_n^{j-1}) \rangle \\ &= \sum_{k=1}^n x_k^{i+j-2} = s_{i+j-2} \end{aligned}$$

are the entries of the Bezoutiant. We have seen in theorem 2.2.1 that in the case of hyperbolic polynomials we have $\sigma(\mathbb{R}^n) = \{z \in \mathbb{R}^n : \tilde{B}(z) \geq 0\}$, where for a real symmetric matrix A let $A \geq 0$ indicate that A is positive semidefinite. (The set $\sigma(V)$ is independent of the choice of generators, see first remark after theorem 6.2.3).

The following theorem provides a generalization of this special case. It is due to Procesi and Schwarz [**33**].

THEOREM 6.2.3. *In the above setting we have $\sigma(V) = \{z \in V /\!\!/ G : \tilde{B}(z) \geq 0\}$.*

SKETCH OF PROOF. Let $W = V \otimes_{\mathbb{R}} \mathbb{C} = V \oplus iV$ be the complexification of V, and consider $\sigma = (\sigma_1, \ldots, \sigma_n)$ as a mapping from W to $\mathbb{C}^n$. Let $K = G^{\mathbb{C}}$ be the unique

complexification of the compact group G. Then G is an analytic subgroup of K, and K is a reductive Lie group, see e.g. [**22**]. We have $\mathbb{C}[W]^K \cong \mathbb{R}[V]^G \otimes_{\mathbb{R}} \mathbb{C}$, and it is generated by the σ_i. Let $\langle .|. \rangle$ denote the K-invariant bilinear forms on W and W^* extending the G-invariant inner products on V and V^*.
Suppose that $z \in V/\!\!/G$. Using results of Kempf-Ness [**16**] and Dadok-Kac [**10**] 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 $K_w = (G_w)^{\mathbb{C}}$ (note that G_w is compact), and $\bar{w} = g.w$ for some $g \in G$, where '$^-$' denotes the complex conjugate. Using Luna's slice theorem (see [**25**]), one can compute $\operatorname{span}_{\mathbb{R}}\{d\sigma_i(w)\}$ and finds that the linear functional $\lambda(w') := \langle w' | iv_2 \rangle$ on W is in this space.
Let us assume now that $z \in V/\!\!/G$ and $\tilde{B}(z) \geq 0$. Use the point $w = v_1 + iv_2$ in W from the previous paragraph. Then we have $0 \leq \tilde{B}(z) = \tilde{B}(\sigma(w)) = B(w) = (\langle d\sigma_i(w) | d\sigma_j(w) \rangle)_{ij}$ which is Gram's matrix of the symmetric bilinear form $\langle .|. \rangle$ on $\operatorname{span}_{\mathbb{R}}\{d\sigma_i(w)\}$. It follows that $\langle .|. \rangle$ is positive semidefinite on $\operatorname{span}_{\mathbb{R}}\{d\sigma_i(w)\}$ and, since $\lambda \in \operatorname{span}_{\mathbb{R}}\{d\sigma_i(w)\}$, we have $\langle \lambda | \lambda \rangle \geq 0$. But

$$\langle \lambda | \lambda \rangle = \langle iv_2 | iv_2 \rangle = -\langle v_2 | v_2 \rangle \leq 0.$$

Consequently, $v_2 = 0$ and $w = v_1 \in V$. So $z = \sigma(v_1) \in \sigma(V)$.
The converse inclusion $\sigma(V) \subseteq \{z \in V/\!\!/G : \tilde{B}(z) \geq 0\}$ is easier: $\sigma(V) \subset V/\!\!/G$ is clear. For arbitrary $v \in V$ consider $\tilde{B}(\sigma(v)) = B(v)$ which is positive semidefinite, since $\langle .|. \rangle$, being an inner product, is so.
A detailed proof can be found in [**33**]. □

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 $\langle .|. \rangle$. Then, relative to these coordinates, B is the matrix of inner products of the gradients of the σ_i; equivalently, $B = JJ^t$, where $J = (\frac{\partial \sigma_i}{\partial v_j})_{ij}$ is the Jacobian of σ. Note that J generalizes the Vandermonde matrix of the symmetric group case (section 2.2).

For each $1 \leq i_1 < \dots < i_s \leq n$ and $1 \leq j_1 < \dots < j_s \leq n$ $(s \leq n)$ consider the matrix with entries $\langle d\sigma_{i_p} | d\sigma_{j_q} \rangle$ for $1 \leq p, q \leq s$, a principal (i.e. symmetric) minor of B. Denote its determinant by $\Delta^{j_1,\dots,j_s}_{i_1,\dots,i_s}$. Then, $\Delta^{j_1,\dots,j_s}_{i_1,\dots,i_s}$ 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 minors $\tilde{\Delta}^{j_1,\dots,j_s}_{i_1,\dots,i_s}(z)$ are non-negative.

6.3. The slice theorem

Let us make a short parenthesis on conjugacy classes, principal orbits and slices.

We consider a Lie group G acting smoothly on a smooth manifold M. We shall write $l : G \times M \to M : (g, x) \mapsto l(g, x) = l_g(x) = l^x(g) = g.x$ for the action and speak of a G-manifold M.

The closed subgroups of G can be partitioned into equivalence classes by the following relation:

$$H \sim H' \quad :\iff \quad \exists\, g \in G \text{ for which } H = gH'g^{-1}.$$

The equivalence class of H is denoted by (H), it is often 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 *isotropy type* of the orbit through x. Two orbits are said to be of the same type, if they have the same isotropy 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:

$$\begin{aligned}(H) \leq (H') \qquad &:\Longleftrightarrow \qquad \exists\ K \in (H) \text{ and } K' \in (H') : K \subseteq K' \\ &\Longleftrightarrow \qquad \exists\ g \in G : H \subseteq gH'g^{-1}.\end{aligned}$$

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 6.3.1. *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 $\bar{A}$.
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 $g^{-1} \in \bar{A}$.
Now suppose that e is discrete 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 $\mathrm{conj} : G \times G \to G, (g,h) \mapsto ghg^{-1}$ is continuous, H is closed and $\mathrm{conj}(A,H) \subseteq H$, as we have seen at the beginning of the proof, we have $\mathrm{conj}(\bar{A},H) \subseteq H$. In particular, $g^{-1}Hg \subseteq H$ which together with our premise implies that $gHg^{-1} = H$. $\square$

DEFINITION 6.3.2. 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 the 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 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.

DEFINITION 6.3.3. 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 following properties of slices:

PROPOSITION 6.3.4. *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 $G_y \subset G_x$, 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 6.3.2), $G_y \subseteq G_x \subseteq gG_yg^{-1}$. Since G_y is compact, $G_y \subseteq gG_yg^{-1}$ implies $G_y = gG_yg^{-1}$, by lemma 6.3.1. Therefore, $G_y = G_x$, and $G.y$ is a principal orbit, too.

To (7): For any $s \in S$ it holds that $(G_x)_s = G_s$: $(G_x)_s \subseteq G_s$ 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. $\square$

The following theorem (due to Koszul even though in a different version) is usually referred to as the *differentiable slice theorem*, since there also exist the *algebraic slice theorem* and the *holomorphic slice theorem*, see [**38**]. 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 $G \to G/G_x$:

$$\begin{tikzcd} G \times S \arrow[r, "q"] \arrow[d, "\mathrm{pr}_1"'] & G \times_{G_x} S \arrow[d, "\bar{p}"] \\ G \arrow[r, "p"] & G/G_x \end{tikzcd}$$

Recall that q is a submersion and $(G \times S, q, G \times_{G_x} S, G_x)$ is a principal bundle.

Theorem 6.3.5. *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:

$$\begin{tikzcd} G \times S \arrow[r, "l"] \arrow[d, two heads, "q"'] & G.S \\ G \times_{G_x} S \arrow[ur, "f"'] & \end{tikzcd}$$

f is smooth because $f \circ q = l$ is smooth and q is a submersion. It is equivariant, since l and q are equivariant:

$$f(h.[g,s]) = f(h.q(g,s)) = f(q(hg,s)) = l(hg,s) = h.l(g,s) = g.f([g,s]),$$

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 6.3.4(2)

$$\begin{aligned} g_1.s_1 = g_2.s_2 \quad &\Leftrightarrow \quad s_1 = g_1^{-1} g_2.s_2 \\ &\Leftrightarrow \quad g_1 = g_2 h^{-1} \text{ and } s_1 = h.s_2 \text{ for } h = g_1^{-1} g_2 \in G_x, \end{aligned}$$

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. □

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 6.3.6 (Existence of slices). [**27**], [**29**]
Let M be a G-space 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. [**27**].

Definition 6.3.7. 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) 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}) and $(g_{n_k}.x_{n_k})$. Now (2) guarantees that we can find a subsequence of (g_{n_k}), and hence of (g_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 $(l, \mathrm{pr}_2)^{-1}(R) \in \{g \in G : g.K \cap L \neq \emptyset\} \times K$. 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. $\square$

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).

THEOREM 6.3.8. *If M is a proper G-manifold, then M/G is completely regular.*

PROOF. Choose $F \in M/G$ closed and $\pi(x_0) \notin F$. 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 (3), $\{g \in G : g.x \in \mathrm{supp} f\}$ 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 (3), the set $\{g \in G : g.V \cap \mathrm{supp} f \neq \emptyset\}$ 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 $\mathrm{supp}\tilde{f} \subseteq G.\mathrm{supp} f \subseteq G.U$, and, consequently, $\mathrm{supp}\tilde{f} \cap \pi^{-1}(F) = \emptyset$. 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 $\bar{f}|_F = 0$. $\square$

Finally, we want to show that the orbits of a proper action are closed submanifolds. For it we need the following lemma:

LEMMA 6.3.9. *A continuous proper map $f : X \to Y$ between two topological spaces 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. Then the $f(a_n)$ are contained in a bounded subset $B \subseteq f(A)$. Therefore, $(a_n) \subseteq f^{-1}(B) \cap A$ which is now, since f is proper, a bounded subset of A. Consequently, (a_n) has a convergent subsequence with limit $a \in A$, and by continuity of f, it gives a convergent subsequence of $(f(a_n))$ with limit $f(a) \in f(A)$. Since $f(a_n)$ converges to y, we have $y = f(a) \in f(A)$. $\square$

PROPOSITION 6.3.10. *The orbits of a proper action $l : G \times M \to M$ are closed submanifolds.*

PROOF. By the preceding lemma 6.3.9, (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 a neighborhood U of e in G, $l^x(U) = U.x$ is a neighborhood of x in $G.x$. Let as assume the contrary: there exits a sequence $(g_n.x) \subseteq G.x \backslash U.x$ which converges to x. Then by definition 6.3.7(2), (g_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. [**21**]. 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. □

6.4. Reducing the problem

We adopt the setting presented in section 6.2. Then, as will turn out after two general definitions, we find slices at each point $v \in V$.

Definition 6.4.1. Let M be a proper Riemannian G-manifold, $x \in M$. The *normal bundle to the orbit* $G.x$ is defined as $\mathrm{Nor}(G.x) := T(G.x)^{\perp}$.
Let $\mathrm{Nor}_\epsilon(G.x) := \{X \in \mathrm{Nor}(G.x) : |X| < \epsilon\}$, 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 $\mathrm{Nor}_{\frac{r}{2}}(G.x)$ onto an open neighborhood of $G.x$, so $\exp(\mathrm{Nor}_{\frac{r}{2}}(G.x)) =: U_{\frac{r}{2}}$ is a tubular neighborhood of $G.x$. We define the *normal slice at* x by $S_x := \exp_x(\mathrm{Nor}_{\frac{r}{2}}(G.x))_x$.

Proposition 6.4.2. *The so defined normal slice* $S_x = \exp_x(\mathrm{Nor}_{\frac{r}{2}}(G.x))_x$ *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:

$$\begin{aligned} S_{g.x} &= \exp_{g.x}(\mathrm{Nor}_{\frac{r}{2}}(G.x))_{g.x} \\ &= \exp_{g.x}(T_x l_g(\mathrm{Nor}_{\frac{r}{2}}(G.x))_x) \\ &= l_g(\exp_x(\mathrm{Nor}_{\frac{r}{2}}(G.x))_x) \\ &= g.S_x, \end{aligned}$$

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,

$$r(l_h(\exp_{g.x} X)) = r(\exp_{hg.x}(T_x l_h.X)) = hg.x = l_h(r(\exp_{g.x} X)),$$

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. □

Definition 6.4.3. Let M be a G-manifold and $x \in M$, then there is a representation of the isotropy group G_x

$$G_x \to GL(T_xM) : g \mapsto T_x l_g$$

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: Observe first 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., $X = \frac{d}{dt}\big|_{t=0} l^x(g_t) \in T_e l^x.\mathfrak{g}$. Consequently, for $g \in G_x$,

$$T_x l_g(T_x(G.x)) = T_x l_g.T_e l^x.\mathfrak{g} = T_e(l_g \circ l^x).\mathfrak{g} = T_e l^x.\mathfrak{g} = T_x(G.x).$$

So $N_x := T_x(G.x)^\perp = \mathrm{Nor}(G.x)_x$ is also G_x-invariant, and

$$G_x \to O(N_x) : g \mapsto T_x l_g$$

is called the *slice representation.*

In our setting, where the compact Lie group G acts orthogonally on 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 6.3.4(8) and theorem 6.3.5, where now we can replace 'smooth' by 'real analytic' anywhere, since the vectorspace 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 question we are interested in it follows:

THEOREM 6.4.4. *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.

Recall the definition of a regular point given in definition 6.3.2. We give now other characterizations in terms of slices and slice representations:

LEMMA 6.4.5. *Let M be a Riemannan 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 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 $G_{g.s} = gG_s g^{-1} = gG_x g^{-1}$. Therefore, $G.x$ is a principal orbit; see remarks after definition 6.3.2. The converse is true by proposition 6.3.4(6), since G_x is compact. □

Let us return to our setting. Assume $v \in V$ is regular. By theorem 6.3.5 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 6.3.4(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 6.3.4(3)).
- To see that V_{reg} is dense in V, let $U \subseteq V$ be open, $x \in U$, and S_x a normal 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 normal 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 6.4.2) for a $g \in G$. Consequently, $G_z \subseteq G_{g.y} = gG_y g^{-1}$ by proposition 6.3.4(5). By choice of y, this implies $G_z = gG_y g^{-1}$ for all $z \in G.S_x \cap G.S_y \cap U$, and $G.y$ is a principal orbit.

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 taken from [**37**].

THEOREM 6.4.6. *In our setting (see section 6.2), for $v \in V$, the subspace $N_v^{G_v}$ of G_v-invariant vectors in N_v is spanned as a real vector space by* $\operatorname{grad}\sigma_1(v), \dots, \operatorname{grad}\sigma_n(v)$.

PROOF. Clearly each $\operatorname{grad}\sigma_i(v) \in N_v^{G_v}$:

$$\langle T_v(G.v)|\operatorname{grad}\sigma_i(v)\rangle = d\sigma_i(v)(T_v(G.v)) = 0,$$

since σ_i is constant on $G.v$, and, for arbitrary $w \in V$ and $g \in G_v$,

$$\begin{aligned}\langle T_v l_g.w|T_v l_g.\operatorname{grad}\sigma_i(v)\rangle &= \langle w|\operatorname{grad}\sigma_i(v)\rangle \\ &= d\sigma_i(v)(w) \\ &= d(\sigma_i \circ l_g)(v)(w) \\ &= d\sigma_i(v)(T_v l_g.w) \\ &= \langle T_v l_g.w|\operatorname{grad}\sigma_i(v)\rangle.\end{aligned}$$

In the following we will identify G with its image $\rho(G) \subseteq O(V)$. 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 6.3.5) 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 6.3.4(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

$$\begin{aligned}|v - x| &= \min_{g\in G}|g.v - x| && \text{since } v = p_{G.v}(x) \\ &= \min_{g\in G}|hg.v - h.x| && \text{for all } h \in G \\ &= \min_{g\in G}|v - g^{-1}.x| && \text{by choosing } h = g^{-1}.\end{aligned}$$

For $w \in G.S_x$ we consider the local, smooth, G-invariant function given by

$$\begin{aligned}\operatorname{dist}(w, G.x)^2 &= \operatorname{dist}(w, p_{G.x}(w))^2 \\ &= \langle w - p_{G.x}(w)|w - p_{G.x}(w)\rangle \\ &= \langle w|w\rangle + \langle p_{G.x}(w)|p_{G.x}(w)\rangle - 2\langle w|p_{G.x}(w)\rangle \\ &= \langle w|w\rangle + \langle x|x\rangle - 2\langle w|p_{G.x}(w)\rangle.\end{aligned}$$

Its derivative with respect to w is

$$d(\operatorname{dist}(\quad, G.x)^2)(w)(y) = 2\langle w|y\rangle - 2\langle y|p_{G.x}(w)\rangle - 2\langle w|dp_{G.x}(w)(y)\rangle. \tag{6.1}$$

We shall show below that

$$\langle v | dp_{G.x}(v)(y) \rangle = 0 \qquad \text{for all } y \in V, \tag{6.2}$$

such that the derivative at v is given by

$$d(\operatorname{dist}(\quad, G.x)^2)(v)(y) = 2\langle v|y\rangle - 2\langle y|p_{G.x}(v)\rangle = 2\langle v - x|y\rangle = -2\langle n|y\rangle. \tag{6.3}$$

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 (since it vanishes outside of $G.S_x$) which coincides with $\operatorname{dist}(\quad, G.x)^2$ near v. By the theorem of Schwarz 6.2.2, there is a smooth function $h \in C^\infty(\mathbb{R}^n, \mathbb{R})$ such that $g = h \circ \sigma$, where $\sigma = (\sigma_1, \dots, \sigma_n) : V \to \mathbb{R}^n$ is the orbit map from section 6.2. Consequently, we have, by equation (6.3),

$$\begin{aligned} -2n &= \operatorname{grad}(\operatorname{dist}(\quad, G.x)^2)(v) \\ &= \operatorname{grad} g(v) = \operatorname{grad}(h \circ \sigma)(v) \\ &= \sum_{i=1}^{n} \frac{\partial h}{\partial y_i}(\sigma(v)) \operatorname{grad} \sigma_i(v), \end{aligned}$$

which proves the result.
It remains to check equation (6.2). We have $T_vV = T_v(G.v) \oplus N_v$, 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 $\langle v|dp_{G.x}(v)(X.v)\rangle = 0$ 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 $g \mapsto \langle v - g.x | v - g.x\rangle$ at e vanishes. Consequently, we have for all $X \in \mathfrak{g}$

$$0 = 2\langle -X.x | v - x\rangle = 2\langle X.x|x\rangle - 2\langle X.x|v\rangle = 0 - 2\langle X.x|v\rangle, \tag{6.4}$$

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 (6.4)

$$\langle v|dp_{G.x}(v)(X.v)\rangle = \langle v|X.x\rangle = 0.$$

This completes the proof. □

6.5. Stratification of the orbit space

This section is dedicated to the study of the natural stratification of the orbit space given by summarizing orbits of the same isotropy type.

Let (H) be one particular orbit type ($H = G_v$ for a $v \in V$), see the beginning of section 6.3. 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 *isotropy stratum of the representation* $\rho : G \to O(V)$, and the image $\sigma(V_H)$ is called an *isotropy stratum of 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 6.3.5. Then 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 6.3.4(5) we have $G_s \subseteq H$. Hence $G_s = H$, by lemma 6.3.1, 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 6.3.4(5), therefore $G_s = H$ and so $[g, s] \in (G \times_H S_v)_H$.

From now on let S_v be the 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 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 proof of theorem 6.3.5) 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. □

In particular, claim 1 yields that V_H is a proper Riemannian G-manifold, since G is compact and since the restriction of $\langle . | . \rangle$ to V_H defines a Riemannian metric on V_H (again denoted by $\langle . | . \rangle$). 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 the normal slice at x (with respect to the action of G on V_H). By proposition 6.3.4(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 6.3.1, they must be the same. That implies, by lemma 6.4.5, that G_x acts trivially on S_x. From proposition 6.3.4(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 6.3.5 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 6.3.8, V_H/G is Hausdorff, and consequently it is a smooth manifold. □

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 $\pi : V_H \to V_H/G$ a Riemannian submersion, i.e., $T_x\pi : \mathrm{Hor}(\pi) := \ker(T\pi)^\perp \to T_{\pi(x)}(V_H/G)$ is an isometric isomorphism for all $x \in V_H$.

Proof. For $X_x, Y_x \in \operatorname{Hor}(\pi)_x = N_x$ we define

$$\gamma_{\pi(x)}(T\pi.X_x, T\pi.Y_x) := \langle X_x | Y_x \rangle_x.$$

This gives a well defined inner product on $T_{\pi(x)}(V_H/G) \cong N_x$: Choose $X'_{g.x}, Y'_{g.x} \in \operatorname{Hor}(\pi)_{g.x}$ such that $T\pi.X'_{g.x} = T\pi.X_x$ and $T\pi.Y'_{g.x} = T\pi.Y_x$ (remember that $T\pi$ is surjective). Then, $T\pi.(X'_{g.x} - Tl_g.X_x) = 0$, so the difference $X'_{g.x} - Tl_g.X_x$ is vertical, i.e., $X'_{g.x} - Tl_g.X_x \in \ker(T\pi)$. On the other hand $X'_{g.x}$ is horizontal, and so is $Tl_g.X_x$, this is because Tl_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, $X'_{g.x} - Tl_g.X_x$ is horizontal as well as vertical and must be zero, i.e., $X'_{g.x} = Tl_g.X_x$, and in the same way $Y'_{g.x} = Tl_g.Y_x$. Now we can conclude that

$$\langle X'_{g.x} | Y'_{g.x} \rangle_{g.x} = \langle Tl_g.X_x | Tl_g.Y_x \rangle_{g.x} = \langle X_x | Y_x \rangle_x.$$

The Riemannian metric γ on V_H/G makes π a Riemannian submersion. □

Let us finally try to understand in what sense $\pi : V_H \to V_H/G$ is an associated bundle. Consider the set $V_H^H := \{x \in V_H : h.x = x \text{ for all } h \in H\}$ of H-invariant points in V_H. We assert that V_H^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 $V_H^H = \bigcap_{h\in H} V_H^h$, and the assertion follows.
Moreover, V_H^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 $x \in V_H^H$. The restriction $\pi : V_H^H \to V_H/G$ is a smooth submersion, since for each $x \in V_H^H$ the corresponding slice S_x is also contained in V_H^H: $G_y = G_x = H$ for all $y \in S_x$ as seen before. The fiber of $\pi : V_H^H \to V_H/G$ 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, $\pi : V_H^H \to V_H/G$ 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 $g.x \in V_H^H$.
So we have proved that $\pi : V_H^H \to V_H/G$ is a principal $N(H)/H$-bundle.

Claim 4. V_H is the associated bundle with fiber G/H.

Proof. Consider the following diagram:

$$\begin{tikzcd}
V_H^H \times G/H \arrow[d] \arrow[dr, "{(x,[g]) \mapsto g.x}"] & \\
V_H^H \times_{N(H)/H} G/H \arrow[r, "\cong"'] \arrow[d] & V_H \arrow[d] \\
V_H/G \arrow[r, equal] & V_H/G
\end{tikzcd}$$

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^H. □

In particular, 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 6.3, namely the principal orbit type. One can prove that from connectedness of V it follows that there is precisely one principal orbit type, see [**7**], [**27**], [**31**]. So $\pi : V_{\text{reg}} \to V_{\text{reg}}/G$ is a locally trivial fiber bundle.

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 for every $x \in V$ 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 $\dim V < m$. Consider the normal slice S_x at x. Then S_x is a Riemannian manifold, and G_x acts isometrically on S_x. By proposition 6.3.4(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. Therefore, it will do to show that the number of G_x-orbit types in S_x is finite. If $\dim S_x < \dim V$, then this follows from the induction hypothesis. Assume $\dim S_x = m$. S_x is an open ball in $N_x = T_xV \cong V$. Since the slice representation is orthogonal, it restricts to a G_x-action on the sphere $S^{m-1} \subseteq V$. By induction hypothesis, locally, S^{m-1} has only finitely many G_x-orbit types. Since S^{m-1} is compact, 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, S_x has only finitely many orbit types: those of S^{m-1} and the 0-orbit.

Hence V and V/G are in a sense stratified, and $\pi : V \to V/G$ is a stratified Riemannian submersion.

We shall show now that the stratification of the orbit space $V/G = \sigma(V)$ by submanifolds $V_H/G = \sigma(V_H)$ 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

$$U^q := W^q \cap \bigcap_{j>q} (\operatorname{int}_q(W^q \cap \bar{\Gamma}^j_\nu) \cup \operatorname{int}_q(W^q \backslash \bar{\Gamma}^j_\nu)),$$

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 [**6**], [**24**]. The following theorem is due to Bierstone [**6**].

THEOREM 6.5.1. *The semianalytic (primary) stratification of the orbit space $\sigma(V)$ coincides with its stratification by components of submanifolds 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 $q = p, p-1, \dots, 0$, by decreasing induction on q.
Obviously, we have $X^p = \sigma(V_{\text{reg}}) = U^p$. 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

the 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 $(y_1, \dots, y_m)$ coordinates in T_v about v, we may choose a set of generators $\psi_1, \dots, \psi_l$ 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 G_v-invariant linear forms on T_v. Let $\psi = (\psi_1, \dots, \psi_l)$ be the corresponding orbit map.
We assert that $\sigma(V)$ is analytically isomorphic near $\sigma(v)$ to a neighborhood of the origin in $S_v^{G_v} \times \psi(T_v) \subseteq S_v^{G_v} \times \mathbb{R}^l$. For: each $\sigma_i \in \mathbb{R}[V]^G$ is, in particular, a G_v-invariant polynomial on $S_v^{G_v} \times T_v$, whence we can write $\sigma_i(x) - \sigma_i(v) = p_i((\mathrm{id} \times \psi)(x - v))$, where $x \in S_v^{G_v} \times T_v$ and p_i is a polynomial function on $S_v^{G_v} \times \mathbb{R}^l$. On the other hand every real analytic G_v-invariant function on $S_v^{G_v} \times T_v$ can be written as a real analytic function in the $\sigma_i(x) - \sigma_i(v)$, see [**39**]. So $P = (p_1, \dots, p_n) : S_v^{G_v} \times \mathbb{R}^l \to \mathbb{R}^n$ provides the required analytic isomorphism of $S_v^{G_v} \times \psi(T_v)$ and $\sigma(V)$ near $\sigma(v)$.
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 $c = (c_1, \dots, c_l)$ 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 7.1.3. 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_ν^j of U^j, $j > q$, we have $\mathrm{int}_q(W^q \backslash \bar{\Gamma}_\nu^j) = W^q \backslash \mathrm{int}_q(W^q \cap \bar{\Gamma}_\nu^j)$:

- ($\subseteq$) Let $z \in \mathrm{int}_q(W^q \backslash \bar{\Gamma}_\nu^j)$. In particular, $z \notin \bar{\Gamma}_\nu^j$ and so $z \notin \mathrm{int}_q(W^q \cap \bar{\Gamma}_\nu^j)$.
- ($\supseteq$) Suppose $z \notin \mathrm{int}_q(W^q \backslash \bar{\Gamma}_\nu^j)$, i.e., each open neighborhood of z in W^q contains accumulation points of $\bar{\Gamma}_\nu^j$. We already know that $X^j = U^j$ for $j > q$, and $X^q = W^q$. Therefore, the piece of boundary of Γ_ν^j lying in W^q, namely $W^q \cap \bar{\Gamma}_\nu^j$, must have dimension q. Consequently, there exists a neighborhood of z in W^q consisting entirely of accumulation points of Γ_ν^j; in other words, $z \in \mathrm{int}_q(W^q \cap \bar{\Gamma}_\nu^j)$.

With this identity we see that $U^q = W^q = X^q$, and the theorem is proved. $\square$

Remark. If the real vector space V is replaced by a complex vector space, then the semianalytic (primary) stratification of the orbit space is coarser than the stratification by orbit type.

As a consequence of theorem 6.4.6, 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:

$$\begin{aligned} \dim V_H/G &= \dim N_v^{G_v} \\ &= \operatorname{rank} d\sigma(v) \\ &= \operatorname{rank} B(v) \\ &= \operatorname{rank} \tilde{B}(\sigma(v)), \end{aligned}$$

where the definition of B and $\tilde{B}$ can be found in section 6.2.

Finally, note that, as seen in the proof of theorem 6.5.1, 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.

To conclude this section let us investigate the stratification of the orbit space in the case when the symmetric group S_n acts on $\mathbb{R}^n$. In this situation we may choose the following fundamental domain

$$F := \{(x_1, \dots, x_n) \in \mathbb{R}^n : x_1 < x_2 < \cdots < x_n\},$$

and we can identify the orbit space $\mathbb{R}^n/S_n$ with the set

$$F \cup \bigcup_{k=2}^{n} (\partial F \cap \{x_1 = x_2 = \cdots = x_k\}).$$

In this picture the principal stratum consisting of all regular orbits equals F, and the stratum of dimension $n - j$, where $1 \leq j \leq n - 1$, is given by

$$(\partial F \cap \{x_1 = \cdots = x_{j+1}\}) \setminus \bigcup_{k>j+1}^{n} \{x_1 = \cdots = x_k\}.$$

CHAPTER 7

Lifting curves over invariants smoothly

This chapter presents many results concerning our lifting problem for real analytic and smooth curves in the orbit space. It is based on [**2**].

7.1. Local lifting

Similarly as in the case of choosing roots of polynomials smoothly, see section 2.3, we shall construct an algorithm which solves the lifting problem locally.

We investigate at first the lifting at regular orbits. This corresponds to lemma 2.3.2. By a *orthogonal lift* we mean a lift meeting orbits orthogonally.

LEMMA 7.1.1. *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 $c(t_0) \in V_{\text{reg}}/G$. It is unique up to a transformation from G.*

PROOF. The orthogonal distribution $V_{\text{reg}} \ni v \mapsto N_v$ of the locally trivial fiber bundle $\pi : V_{\text{reg}} \to V_{\text{reg}}/G$ 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 [**21**], section 9. $\square$

To lift a curve in the neighborhood of a singular orbit is more involved. We shall need two lemmas. First we learn how to deal with nontrivial 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 7.1.2. *Any lift $\bar{c}$ of a curve $c = (c_0, c_1)$ of class C^k $(k = 0, 1, \ldots, \infty, \omega)$ 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 $(k = 0, 1, \ldots, \infty, \omega)$. The lift $\bar{c}$ is orthogonal if and only if the lift $\bar{c}_1$ is orthogonal.*

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 2.3.4:

$$m(f) := \sup\{p \in \mathbb{Z} : f(t) = t^p g(t) \text{ near } 0 \text{ for continuous } g\}.$$

Let $c = (c_1, \ldots, c_n)$ be a smooth curve in $\sigma(V) \subseteq \mathbb{R}^n$ with $c(0) = 0$. By possibly increasing the number of generators $\sigma_1, \ldots, \sigma_n$ of $\mathbb{R}[V]^G$, we may assume from now on without loss that $\sigma_1 : v \mapsto \langle v|v \rangle$ 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 7.1.3 (Multiplicity lemma). *In this situation we have $m(c_i) \geq d_i r$, for all $1 \leq i \leq n$. Remember $d_1, \ldots, d_n$ are the degrees of homogeneity of the generators $\sigma_1, \ldots, \sigma_n$.*

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 6.2.3, 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. □

If $r < \infty$, we shall consider the following smooth curve in $\sigma(V)$:

$$c_{(r)}(t) := (t^{-2r}c_1(t), t^{-d_2 r}c_2(t), \dots, t^{-d_n r}c_n(t)). \tag{7.1}$$

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.

After this preliminary work we can attack the local lifting problem for real analytic curves.

THEOREM 7.1.4. *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 2.3.9.

(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 7.1.1.
(2) If $V^G \neq \{0\}$, then we remove fixed points, by lemma 7.1.2. 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 $G_v \to O(N_v)$. By theorem 6.4.4, 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 $c_1 \equiv 0$, since c_1 is real analytic. This implies that $c \equiv 0$ is constant which clearly can be lifted. In the former case, by the multiplicity lemma 7.1.3, 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 (7.1). Then $c_{(r)}(0) \neq 0$, and we may continue in 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 2.3.3 in part 1 is now played by the transition to the slice representation provided by theorem 6.4.4.
(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 as the 'most' singular point.

The forgoing theorem 7.1.4 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 2.3 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 7.1.4 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 conditions for the smooth curve c in the orbit space, we have to take care that this implication remains valid.

Definition 7.1.5. Let $s \in \mathbb{N}_0$. Denote by A_s the union of all strata X of the orbit space V/G with $\dim X \le s$, 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 7.1.4.

Proposition 7.1.6. *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 the slice representation or replacing c by $c_{(r)}$ (see equation* (7.1)*), 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 at the beginning of the section, 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 the slice representation: 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 6.5) and from theorem 6.4.6, 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. $\square$

The following theorem gives the best practical way to check the normal nonflatness of a curve c, in terms of the principal minors $\tilde{\Delta}^{j_1,\dots,j_k}_{i_1,\dots,i_k}$ of $\tilde{B}$, see section 6.2.

THEOREM 7.1.7. *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 $1 \le r \le n$:*

(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 $s = s(c, t_0)$ again 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 section 6.5, condition (1) yields that $s \le r$. Condition (2) guarantees that $s = r$. 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. $\square$

With these ingredients we can attack the problem of lifting smooth curves locally:

THEOREM 7.1.8. *Let $c = (c_1, \dots, c_n) : \mathbb{R} \to \sigma(V) \subseteq \mathbb{R}^n$ 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 7.1.4, since by proposition 7.1.6 the normal nonflatness remains invariant under the reduction process and it guarantees that in step (4) from $r = \infty$ follows $c_1 \equiv 0$. $\square$

Let us conclude this section with a result concerning the uniqueness of local lifts:

LEMMA 7.1.9. *Let $c : \mathbb{R} \to \sigma(V) \subseteq \mathbb{R}^n$ 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 7.1.4.
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 the 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 7.1.2. 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 (7.1). 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. □

7.2. Global lifting

Here we shall glue together the local smooth lifts found in the previous section in order to get a global smooth lift.

THEOREM 7.2.1. *Let $c : \mathbb{R} \to \sigma(V) \subseteq \mathbb{R}^n$ be a generic smooth curve. Then there exists a global smooth lift $\bar{c} : \mathbb{R} \to V$ with $\sigma \circ \bar{c} = c$.*

PROOF. By theorem 7.1.8, 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 7.1.8, 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 7.1.9 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 $t \leq t_1$ 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,

$$\bar{c}_{12}(t) := \begin{cases} \bar{c}_1(t) & \text{for} \quad t \leq t_1 \\ h(t).\bar{c}_2(t) & \text{for} \quad t \geq t_1 \end{cases}$$

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-partitions of unity.

7.3. Polar representations

In this section we show that, if we restrict to a smaller class 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 intend a lift meeting orbits orthogonally.

The mentioned smaller class of representations is the one of polar representations:

DEFINITION 7.3.1. An orthogonal representation $\rho : G \to O(V)$ 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 [**9**], [**10**] and [**30**].

Suppose we are given a polar representation $\rho : G \to O(V)$ 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. $Z(\Sigma) = \bigcap_{s\in\Sigma} G_s$ is closed as well, and it is a normal subgroup of $N(\Sigma)$. 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

$$W(\Sigma) = N(\Sigma)/Z(\Sigma).$$

$W(\Sigma)$ is a finite group: Take a regular point $v \in \Sigma$ and consider the normal slice S_v at v. Then $S_v \subseteq \Sigma$ open. Hence, any $g \in N(\Sigma)$ close to the identity element maps v back into S_v. By proposition 6.3.4(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.
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 7.3.2. [**10**], [**40**]
Let $\rho : G \to O(V)$ 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, including their stratifications.

THEOREM 7.3.3. *Let $\rho : G \to O(V)$ 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 $c : \mathbb{R} \to \sigma(V) \subseteq \mathbb{R}^n$ 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 7.3.2, $\sigma|_\Sigma : \Sigma \to \mathbb{R}^n$ is the orbit map for the representation $W(\Sigma) \to O(\Sigma)$. If c is a generic smooth curve in $\sigma(V) \cong \sigma|_\Sigma(\Sigma)$, then by theorem 7.2.1 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 7.1.7.
If c is real analytic, there are local real analytic lifts over $\sigma|_\Sigma$ into Σ by theorem 7.1.4. By lemma 7.1.9, 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 in the first two paragraphs of the proof. Since c is generic (remember a real analytic curve is automatically 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 6.3.4(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 6.3.1), call them H, and we have $H \subseteq G_{\bar{c}_1(t)}$ for all $t \in \mathbb{R}$.
From lemma 7.1.9 we get that $\bar{c}_1(t) = g(t).\bar{c}_2(t)$ for some smooth or real analytic

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_0. Consequently, we find $\bar{c}_1(t) = g.\bar{c}_2(t)$ for all $t \in \mathbb{R}$. $\square$

CHAPTER 8

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 $\sigma(V)$. In the smooth case we saw that one has to impose certain genericity conditions on c, see definition 7.1.5, 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 the example at the beginning of section 5.1, 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 [**19**].

8.1. Lifting curves continuously

In this section we shall lift curves continuously.

THEOREM 8.1.1. *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 make 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 $G' \to O(V)$ 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 7.1.2, 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, acting orthogonally on N_v, satisfy $G_v < G$. Therefore, by induction hypothesis and by theorem 6.4.4, 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,

$$\bar{c}_{12}(t) := \begin{cases} \bar{c}_1(t) & \text{for} \quad t \leq t_0 \\ g.\bar{c}_2(t) & \text{for} \quad t \geq t_0 \end{cases}$$

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.

Now let us observe what happens with the lift $\bar{c} : (a_i,b_i) \to V\backslash\{0\}$ of $c|_{(a_i,b_i)}$ constructed in the previous paragraph, if $t \searrow a_i \neq -\infty$ or $t \nearrow b_i \neq +\infty$. The two cases are completely analogous, so let us assume that $t \searrow a_i \neq -\infty$. Evidently, it implies $c(t) \to 0$. We put $\bar{c}(a_i) := 0$, since, by $\sigma^{-1}(0) = \{0\}$, this is the only choice. What remains to show is that $\bar{c}(t)$ converges to 0 as $t \searrow a_i$. But this is obvious, since $\sigma_1(\bar{c}(t)) = \langle \bar{c}(t)|\bar{c}(t)\rangle = c_1(t) \to 0$ as $t \searrow a_i$. Hence, we have shown that the continuous lifts already found on the open intervals (a_i,b_i) can be extended continuously to their closure.

For isolated points of $c^{-1}(0)$ we can simply put together the two lifts on the neighboring intervals, and we obtain a continuous lift on their union. So we have extended our continuous lift $\bar{c}$ on $\mathbb{R}\backslash E$, where E is the set of accumulation points of $c^{-1}(0)$. Let t' be an accumulation point of $c^{-1}(0)$, i.e., $t' \in E$. Since $c^{-1}(0)$ is closed, we have $c(t') = 0$, and thus every lift of c has to vanish at t'. So let us extend our continuous lift $\bar{c}$ on $\mathbb{R}\backslash E$ to E, by putting $\bar{c}(t) := 0$ for all $t \in E$. To finish the proof we have to show that then $\bar{c}$ is continuous at any $t' \in E$. Again we may consider $\sigma_1(\bar{c}(t)) = \langle \bar{c}(t)|\bar{c}(t)\rangle = c_1(t) \to 0$ as $t \to t'$, which gives the required continuity at t'. Therefore, we obtain a global continuous lift of c. This completes the induction and thus the proof. □

Remark. Note that in proposition 2.4.1 we constructed a continuous lift $x = (x_1,\dots,x_n)$ of the continuous curve $P : \mathbb{R} \to \sigma(\mathbb{R}^n) = \mathbb{R}^n/S_n$ which, moreover, lay in the closure of the fundamental domain

$$F = \{(y_1,\dots,y_n) \in \mathbb{R}^n : y_1 < y_2 < \dots < y_n\}.$$

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 $\bar{F}$.

Nevertheless, one can prove 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.

Theorem 8.1.1 is true also in a more general setting, namely if we replace the vector space V by an arbitrary G-space X, where G is still a compact Lie group. The first step in proving it is the following lemma. This approach is due to Montgomery and Yang [**28**].

Lemma 8.1.2. *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. One can show that in this situation there are slices at every point of X, see [**7**].

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 $\tau(t) := g_0 g_1 \cdots g_i.\tau_i(t)$ 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, making 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 G and its image $F^* := \pi(F) \subseteq I = X/G$ under π, which is closed in I since F is closed, see lemma 8.1.3. 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_α) 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 lemma 8.1.3. $\square$

LEMMA 8.1.3. *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. $\square$

Now we can consider the general case.

THEOREM 8.1.4. *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 $\bar{c} : I \to X$.*

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 lemma 8.1.2, 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. $\square$

8.2. 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'.
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 7.1.1. Recall that by $d_1, \ldots, d_n$ we have denoted the degrees of the homogeneous generators $\sigma_1, \ldots, \sigma_n$, and that $d_1 = 2$ by our choice of $\sigma_1 : v \mapsto \langle v|v \rangle$.

LEMMA 8.2.1. *A curve $c : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ of class C^d, where $d := \max\{d_1, \ldots, d_n\}$ (≥ 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.*

Moreover, we shall need the following stronger version of the multiplicity lemma 7.1.3:

LEMMA 8.2.2. *Let $c = (c_1, \ldots, c_n)$ be a curve in $\sigma(V) \subseteq \mathbb{R}^n$, where c_i is C^{d_i}, for $1 \leq i \leq n$, 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 $1 \leq i \leq n$.*

PROOF. The proof of the nontrivial implication (1) $\Rightarrow$ (2) is the same as in the smooth case with $r = 1$, see the proof of lemma 7.1.3. 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 $c_{(m)}(t) = (t^{-d_1 m} c_1(t), \ldots, t^{-d_n m})$ is continuous for $t \geq 0$. $\square$

The following proposition shows that we can lift a C^d-curve in the orbit space locally near every point $t_0 \in \mathbb{R}$ in such a way that the lift is differentiable in the point t_0.

PROPOSITION 8.2.3. *Let $c = (c_1, \ldots, c_n) : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a curve of class C^d, where $d = \max\{d_1, \ldots, d_n\}$. 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 7.1.4. 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 8.2.1.
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 $c_{1,1}$, since c_1 is C^2. By the variant of the multiplicity lemma 8.2.2, we find that $c_i(t) = t^{d_i} c_{i,i}(t)$ near 0 for $1 \leq i \leq n$, where $c_{1,1}, c_{2,2}, \ldots, c_{n,n}$ are continuous functions. We consider the following continuous curve in $\sigma(V)$

$$\begin{aligned} c_{(1)}(t) &:= (c_{1,1}(t), c_{2,2}(t), \ldots, c_{n,n}(t)) \\ &= (t^{-2} c_1(t), t^{-d_2} c_2(t), \ldots, t^{-d_n} c_n(t)). \end{aligned}$$

By theorem 8.1.1, 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 $G_v \to O(N_v)$. By theorem 6.4.4, the lifting problem reduces to the same problem for C^d-curves in N_v/G_v now passing through 0. Note that the restrictions $\sigma_1|_{N_v}, \dots, \sigma_n|_{N_v}$ generate the algebra of germs at $0 \in N_v$ of G_v-invariant analytic functions on N_v. In particular, each G_v-invariant polynomial on N_v is an analytic function of $\sigma_1|_{N_v}, \dots, \sigma_n|_{N_v}$ near 0. Consequently, we can refer to the previous case, and the theorem is proved. □

8.3. 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. Throughout the whole section we put $d := \max\{d_1, ..., d_n\}$, where $d_i = \deg \sigma_i$. Recall that we assumed $\sigma_1 : v \mapsto \langle v|v \rangle$ and therefore $d_1 = 2$.

At first let us consider a lemma of topological nature.

LEMMA 8.3.1. *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 $F \cap A = \emptyset$. □

Now we can prove the existence of a global differentiable lift.

THEOREM 8.3.2. *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. Then there exists a global differentiable lift $\bar{c} : \mathbb{R} \to V$ of c.*

PROOF. The proof, as the one of theorem 8.1.1, 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 $c : \mathbb{R} \to V/G'$, satisfying the differentiability conditions of the theorem, there exists a global differentiable lift $\bar{c} : \mathbb{R} \to V$ of c, where $G' \to O(V)$ is an orthogonal representation on an arbitrary real finite dimensional Euclidean vector space V.
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 7.1.2. As in the proof of theorem 8.1.1 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, acting on N_v, satisfy $G_v < G$. Therefore, by induction hypothesis and by theorem 6.4.4, 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. For this purpose we split up the derivatives $\bar{c}_1'(t_0)$ and $\bar{c}_2'(t_0)$ in their $T_{v_0}(G.v_0)$-part and in their N_{v_0}-part:

$$\bar{c}_i'(t_0) = \mathrm{pr}_{T_{v_0}(G.v_0)}(\bar{c}_i'(t_0)) + \mathrm{pr}_{N_{v_0}}(\bar{c}_i'(t_0)) \qquad i = 1, 2.$$

Let us first deal with the respective N_{v_0}-parts. We consider the following 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, $t \mapsto g_i(t)^{-1}.\bar{c}_i(t)$ are differentiable lifts of $c|_{(a_i,b_i)}$ near t_0 which lie in S_{v_0}, whence $\frac{d}{dt}\big|_{t=t_0} (g_i(t)^{-1}.\bar{c}_i(t)) = -g_i'(t_0).v_0 + \bar{c}_i'(t_0) \in N_{v_0}$. Hence,

$$\begin{aligned} 0 &= \mathrm{pr}_{T_{v_0}(G.v_0)}\left(\frac{d}{dt}\bigg|_{t=t_0} (g_i(t)^{-1}.\bar{c}_i(t))\right) \\ &= \mathrm{pr}_{T_{v_0}(G.v_0)}(-g_i'(t_0).v_0) + \mathrm{pr}_{T_{v_0}(G.v_0)}(\bar{c}_i'(t_0)) \\ &= -\underbrace{T_e l^{v_0}.g'(t_0)}_{\in T_{v_0}(G.v_0)} + \mathrm{pr}_{T_{v_0}(G.v_0)}(\bar{c}_i'(t_0)). \end{aligned}$$

So, $\mathrm{pr}_{T_{v_0}(G.v_0)}(\bar{c}_i'(t_0)) = T_e l^{v_0}.g'(t_0) = g_i'(t_0).v_0$, and for the N_{v_0}-part of $\bar{c}_i'(t_0)$ we find

$$\mathrm{pr}_{N_{v_0}}(\bar{c}_i'(t_0)) = \frac{d}{dt}\bigg|_{t=t_0} (g_i(t)^{-1}.\bar{c}_i(t)). \tag{8.1}$$

That means that, in order to make match the N_{v_0}-parts of $\bar{c}_1'(t_0)$ and $\bar{c}_2'(t_0)$, we may deal with the differentiable lifts $\tilde{c}_1(t) := g_1(t)^{-1}.\bar{c}_1(t)$ and $\tilde{c}_2(t) := g_2(t)^{-1}.\bar{c}_2(t)$ (instead of $\bar{c}_1$ and $\bar{c}_2$) which lie in S_{v_0} for t close to t_0, and, hence, $\tilde{c}_1'(t_0)$ and $\tilde{c}_2'(t_0)$ lie in N_{v_0}. Now we can change to the isotropy representation $G_{v_0} \to O(N_{v_0})$, and we can suppose that $v_0 = 0$, i.e., $c(t_0) = 0$ and $\tilde{c}_1(t_0) = \tilde{c}_2(t_0) = 0$. Let us remind of the continuous curve in the orbit space defined in the proof of proposition 8.2.3:

$$c_{(1,t_0)}(t) := ((t - t_0)^{-2} c_1(t), (t - t_0)^{-d_2} c_2(t), \ldots, (t - t_0)^{-d_n} c_n(t)).$$

Although we used the curve $c_{(1,t_0)}$ there only in the singular case, it makes perfectly sense also, if $c(t_0)$ is regular. Note that it is depending on the point t_0. 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 $\tilde{c}_1'(t_0) = g_0.\tilde{c}_2'(t_0)$, i.e., $\mathrm{pr}_{N_{v_0}}(\bar{c}_1'(t_0)) = g_0.\mathrm{pr}_{N_{v_0}}(\bar{c}_2'(t_0))$. We will show soon that for $g_0 \in G_{v_0}$

$$g_0.\mathrm{pr}_{N_{v_0}}(\bar{c}_2'(t_0)) = \mathrm{pr}_{N_{v_0}}(g_0.\bar{c}_2'(t_0)). \tag{8.2}$$

So we have already achieved that the N_{v_0}-parts of the derivatives at t_0 of $\bar{c}_1$ and $g_0.\bar{c}_2$ match. We still have $\bar{c}_1(t_0) = g_0.\bar{c}_2(t_0) = v_0$, since $g_0 \in G_{v_0}$.
Now let us concentrate on the $T_{v_0}(G.v_0)$-parts of the derivatives $\bar{c}_1'(t_0)$ and $g_0.\bar{c}_2'(t_0)$. We search for a differentiable curve $t \mapsto g(t)$ in G with $g(t_0) = e$ and with

$$\begin{aligned}
\mathrm{pr}_{T_{v_0}(G.v_0)}(\bar{c}_1'(t_0)) &= \mathrm{pr}_{T_{v_0}(G.v_0)}\left(\left.\frac{d}{dt}\right|_{t=t_0}(g(t)g_0.\bar{c}_2(t))\right)\\
&= \mathrm{pr}_{T_{v_0}(G.v_0)}(g'(t_0).v_0 + g_0.\bar{c}_2'(t_0))\\
&= \underbrace{T_e l^{v_0}.g'(t_0)}_{\in T_{v_0}(G.v_0)} + \mathrm{pr}_{T_{v_0}(G.v_0)}(g_0.\bar{c}_2'(t_0)).
\end{aligned}$$

But this linear equation can be solved for $g'(t_0)$, since the map $T_e l^{v_0}$ is onto $T_{v_0}(G.v_0)$, and, hence, the required curve $t \mapsto g(t)$ exists. So we may replace the differentiable lift $g_0.\bar{c}_2$ by the differentiable lift $t \mapsto g(t)g_0.\bar{c}_2(t)$ in order to make match the $T_{v_0}(G.v_0)$-parts of the corresponding derivatives at t_0. The N_{v_0}-part of $g_0.\bar{c}_2'(t_0)$ remains unchanged by this replacement:

$$\mathrm{pr}_{N_{v_0}}\left(\left.\frac{d}{dt}\right|_{t=t_0}(g(t)g_0.\bar{c}_2(t))\right) = \mathrm{pr}_{N_{v_0}}(g'(t_0).v_0) + \mathrm{pr}_{N_{v_0}}(g_0.\bar{c}_2'(t_0)) = \mathrm{pr}_{N_{v_0}}(g_0.\bar{c}_2'(t_0)),$$

since $g'(t_0).v_0 \in T_{v_0}(G.v_0)$.
Summarizing what we did so far:

$$\bar{c}_{12}(t) := \begin{cases} \bar{c}_1(t) & \text{for} \quad t \le t_0 \\ g(t)g_0.\bar{c}_2(t) & \text{for} \quad t \ge t_0 \end{cases}$$

is a local differentiable lift of $c|_{(a_i,b_i)}$ defined on a larger interval than $\bar{c}_1$. Consequently, we have shown that each local differentiable lift of $c|_{(a_i,b_i)}$ defined on $(a,b) \subseteq (a_i,b_i)$ can be extended to a larger interval whenever $(a,b) \subsetneq (a_i,b_i)$. This proves the claim.

For the proof of (8.2) let $t \mapsto \hat{g}_2(t)$ be a differentiable lift to G, with $\hat{g}_2(t_0) = e$, of the curve $t \mapsto p(g_0.\bar{c}_2(t)) \in G/G_{v_0}$, related to the principal bundle $\pi : G \to G/G_{v_0}$. Then we have (using the identification $G/G_{v_0} \cong G.v_0$)

$$\hat{g}_2(t).v_0 = \pi(\hat{g}_2(t)) = p(g_0.\bar{c}_2(t)) = g_0.p(\bar{c}_2(t)) = g_0.\pi(g_2(t)) = g_0 g_2(t).v_0. \tag{8.3}$$

As above we find

$$\begin{aligned}
\mathrm{pr}_{N_{v_0}}(g_0.\bar{c}_2'(t_0)) &= \left.\frac{d}{dt}\right|_{t=t_0}(\hat{g}_2(t)^{-1}g_0.\bar{c}_2(t)) \qquad \text{(see (8.1))}\\
&= -\hat{g}_2'(t_0).v_0 + g_0.\bar{c}_2'(t_0) \qquad (\text{since } g_0 \in G_{v_0})\\
&= -g_0 g_2'(t_0).v_0 + g_0.\bar{c}_2'(t_0) \qquad \text{(by (8.3))}\\
&= g_0.\left(\left.\frac{d}{dt}\right|_{t=t_0}(g_2(t)^{-1}.\bar{c}_2(t))\right)\\
&= g_0.\mathrm{pr}_{N_{v_0}}(\bar{c}_2'(t_0)),
\end{aligned}$$

and (8.2) is established.

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 $\bar{c}(a_i) := 0$, being the only choice, since $c(a_i) = 0$. 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 gain that the limit $\lim_{t\searrow a_i}\gamma(t)$ exists, otherwise put, that the one-sided derivative of $\bar{c}$ at a_i exists. For t sufficiently close to a_i we have

$$\sigma(\gamma(t)) = \sigma\left(\frac{\bar{c}(t)}{t-a_i}\right) = c_{(1,a_i)}(t) \to c_{(1,a_i)}(a_i) \qquad \text{as } t \searrow a_i, \tag{8.4}$$

where now $c_{(1,a_i)}(t) := ((t-a_i)^{-2}c_1(t),(t-a_i)^{-d_2}c_2(t),\dots,(t-a_i)^{-d_n}c_n(t))$, a continuous curve in $\sigma(V)$. Let $\bar{c}_{(1,a_i)}$ be a continuous lift of $c_{(1,a_i)}$ which exists by theorem 8.1.1. Then (8.4) 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 8.3.1 gives that A is connected (we can find a closed ball containing $\gamma(t)$ for t close to a_i). In particular, the limit $\lim_{t\searrow a_i}\gamma(t)$ must exist, if G is a finite group. In the general situation 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 $t>a_i$, $t\mapsto\gamma(t)$ is a differentiable curve 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 (via the horizontal lift of the principal connection, say). We may suppose that the differentiable curve $t\mapsto g(t)$ is defined on the whole interval (a_i,b_i) and that it becomes identically e for t outside a small neighborhood of a_i (note that $t\mapsto g(t)$ lies in the connected component of e in G, since v_1 is an accumulation point of $(\gamma(t))_{t\searrow a_i}$). 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 $G.v_1\cap S_{v_1}=\{v_1\}$, by (8.4) which we can apply again, 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 $c|_{(a_i,b_i)}$ on the whole interval (a_i,b_i) whose one-sided derivative at a_i exists:

$$\lim_{t\searrow a_i}\frac{g(t)^{-1}.\bar{c}(t)}{t-a_i}=\lim_{t\searrow a_i}g(t)^{-1}.\gamma(t)=v_1.$$

By proposition 8.2.3, there is a local lift of c near a_i which is differentiable at a_i with derivative $\bar{c}_{(1,a_i)}(a_i)$ (analyze the proof of theorem 8.2.3). Since $v_1\in A\subseteq G.\bar{c}_{(1,a_i)}(a_i)$, there is a $g\in G$ such that $v_1=g.\bar{c}_{(1,a_i)}(a_i)$.

Now take the differentiable lift $t\mapsto g(t)^{-1}.\bar{c}(t)$ and apply the same reasoning for $b_i\neq+\infty$. Thus we obtain a differentiable lift of c on 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 (closed) intervals constructed above can be easily made 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 (8.4), where $\bar{c}_{(1,t_0)}$ is a continuous lift of the curve $c_{(1,t_0)}(t) := ((t-t_0)^{-2}c_1(t),\dots,(t-t_0)^{-d_n}c_n(t))$, see theorem 8.1.1. 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 $\langle\tilde{c}(t)|\tilde{c}(t)\rangle=\sigma_1(\tilde{c}(t))=c_1(t)\to 0$ 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 $t\mapsto\frac{\tilde{c}}{t-t'}$ 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 8.2.3. 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 $c_{(1,t')}(t')=\lim_{t\to t'}\sigma(\frac{\bar{c}(t)}{t-t'})=\sigma(\bar{c}'(t'))=\sigma(0)=0$. From this we see that $\lim_{t\to t'}\frac{\tilde{c}}{t-t'}=0$, since $\langle\frac{\tilde{c}(t)}{t-t'}|\frac{\tilde{c}(t)}{t-t'}\rangle=\sigma_1(\frac{\tilde{c}(t)}{t-t'})=(t-t')^{-2}c_1(t)$, the first component of $c_{(1,t')}$.

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. $\square$

REMARK. Note that, if in the choice of the generators $\sigma_1, \dots, \sigma_n$ of $\mathbb{R}[V]^G$ we require additionally that the degrees $d_1, \dots, d_n$ are as less as possible, then the differentiability conditions of the curve c in the current section 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 $1, \dots, n$, there must not exist a differentiable lift, if the differentiability assumptions made on c are weakened, see the first example after lemma 2.1.2.

8.4. Global orthogonal differentiable lifts of smooth curves

We shall prove in this section the existence of a global orthogonal differentiable lift of a smooth curve in the orbit space. But we have to make a few considerations first.

LEMMA 8.4.1. *Let $v \in V$ and $S_v \subseteq N_v$ the normal slice at v. For each $s \in S_v$ we can decompose the Lie algebra $\mathfrak{g} := \mathrm{Lie}(G)$ of G as follows*

$$\mathfrak{g} = \mathfrak{g}^\top(S_v, s) + \mathfrak{g}^\perp(S_v, s),$$

where $\mathfrak{g}^\top(S_v, s) := \{X \in \mathfrak{g} : \zeta_X(s) \in N_v\}$ and $\mathfrak{g}^\perp(S_v, s) := \{X \in \mathfrak{g} : \zeta_X(s) \perp N_v\}$, and ζ_X is the fundamental vector field. Then, for all $s \in S_v$ we have

$$\mathfrak{g}^\top(S_v, s) = \mathfrak{g}_v,$$

where $\mathfrak{g}_v := \mathrm{Lie}(G_v)$ denotes the isotropy subalgebra.

PROOF. Clearly, $\mathfrak{g}_v \subseteq \mathfrak{g}^\top(S_v, s)$: Let $X \in \mathfrak{g}_v$ which means that there is a smooth curve $t \mapsto g(t)$ in G_v with $g(0) = e$ such that $X = \frac{d}{dt}\big|_{t=0} g(t)$ and, hence, $\zeta_X(s) = T_e l^s . \frac{d}{dt}\big|_{t=0} g(t) = \frac{d}{dt}\big|_{t=0} l^s(g(t)) \in N_v$, since $t \mapsto l^s(g(t)) = g(t).s$ is a curve in S_v.
On the other hand, suppose that $X \in \mathfrak{g}^\top(S_v, s)$. Then, $\zeta_X(s) = \frac{d}{dt}\big|_{t=0} (g(t).s) \in N_v$, where $X = \frac{d}{dt}\big|_{t=0} g(t)$ for a smooth curve $t \mapsto g(t)$ in G with $g(0) = e$. Consider the projection $p : G.S_v \cong G \times_{G_v} S_v \to G/G_v \cong G.v$ of a fiber bundle associated to the principal bundle $\pi : G \to G/G_v$. Then, $t \mapsto p(g(t).s)$ is a smooth curve in G/G_v which admits a smooth lift $t \mapsto h(t)$ into G with $h(0) = e$ (via the horizontal lift of the principal connection, say). So, $t \mapsto h(t)^{-1}g(t).s$ defines a smooth curve in S_v, and, consequently, the smooth curve $t \mapsto h(t)^{-1}g(t)$ lies in G_v (since $g.S_v \cap S_v \neq \emptyset$ implies $g \in G_v$). Therefore, $\frac{d}{dt}\big|_{t=0} h(t)^{-1}g(t) = -h'(0) + X \in \mathfrak{g}_v$. Moreover, if we differentiate with respect to t at 0 the equation $p(g(t).s) = \pi(h(t))$, we get $T_s p . T_e l^s . X = T_e \pi . h'(0)$, whose left-hand side is 0, since $T_e l^s . X \in N_v$. But then, $h'(0)$ has to be in $\mathfrak{g}_v$ and, therefore, $X \in \mathfrak{g}_v$. $\square$

As a consequence we can improve theorem 7.1.8 to provide even an orthogonal smooth lift in neighborhoods of normally nonflat points:

COROLLARY 8.4.2. *Let $c = (c_1, \dots, c_n) : \mathbb{R} \to \sigma(V) \subseteq \mathbb{R}^n$ be a smooth curve which is normally nonflat at $t_0 \in \mathbb{R}$. Then there exists an orthogonal smooth lift $\bar{c}$ in V of c, locally near t_0.*

PROOF. The proof is the same as the one of theorem 7.1.8. We only have to check additionally that orthogonality is invariant under the used reduction process: For the reduction by removing fixed points and by replacing c by $c_{(r)}$ this has already been remarked in lemma 7.1.2 and shortly after lemma 7.1.3.
Passing to the slice representation: Suppose $\bar{c} : I \to S_v$ is a local smooth lift of c passing through $v \in V$ which lies in the normal slice S_v at v and is orthogonal to

all G_v-orbits it meets. We have to show that $\bar{c}$ is orthogonal also as curve in V with respect to the G-orbits. Decompose the tangent space of the G-orbit through $\bar{c}(t)$ for each $t \in I$

$$\begin{aligned} T_{\bar{c}(t)}(G.\bar{c}(t)) &= \{\zeta_X(\bar{c}(t)) : X \in \mathfrak{g} = \mathrm{Lie}(G)\} \\ &= \{\zeta_X(\bar{c}(t)) : X \in \mathfrak{g}^\top(S_v, \bar{c}(t))\} + \{\zeta_X(\bar{c}(t)) : X \in \mathfrak{g}^\perp(S_v, \bar{c}(t))\} \\ &= \{\zeta_X(\bar{c}(t)) : X \in \mathfrak{g}_v\} + \{\zeta_X(\bar{c}(t)) : X \in \mathfrak{g}^\perp(S_v, \bar{c}(t))\}, \end{aligned}$$

according to the decomposition of $\mathfrak{g}$ in lemma 8.4.1. Then, we have that $\bar{c}'(t)$ is orthogonal to the first summand, since the lift $\bar{c}$ is orthogonal to the G_v-orbits in S_v, and orthogonal to the second summand, since $\bar{c}$ lies in S_v and so $\bar{c}'(t) \in N_v$. This completes the proof. □

Now we can prove the promised theorem.

THEOREM 8.4.3. *Let $c = (c_1, \ldots, c_n) : \mathbb{R} \to V/G = \sigma(V) \subseteq \mathbb{R}^n$ be a smooth curve. Then there exists a global orthogonal differentiable lift $\bar{c} : \mathbb{R} \to V$ of c.*

PROOF. Let us repeat the proof of theorem 8.3.2, taking care about orthogonality and making a few slight changes. Again we make induction on the size of G. If $G = \{e\}$ is trivial, then $\bar{c} := c$ is a global orthogonal differentiable lift, since each orbit $G.x$ consists of one point $\{x\}$ only and so the whole vector space V is orthogonal to $G.x$.

So let us assume that for any $G' < G$ and any smooth $c : \mathbb{R} \to V/G'$ there exists a global orthogonal differentiable lift $\bar{c} : \mathbb{R} \to V$ of c, where $G' \to O(V)$ is an orthogonal representation on an arbitrary real finite dimensional Euclidean vector space V.

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 smooth. We may assume that $V^G = \{0\}$, by lemma 7.1.2. 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 an orthogonal 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 acting on N_v satisfy $G_v < G$. Therefore, by induction hypothesis and by theorem 6.4.4, we find local differentiable lifts $\bar{c} : I(t) \to S_v \subseteq V$ of $c|_{(a_i,b_i)}$ defined on open neighborhoods $I(t) \subseteq (a_i, b_i)$ of any $t \in (a_i, b_i)$ and through all $v \in \sigma^{-1}(c(t))$ which are orthogonal with respect to the G_v-orbits in S_v. By exactly the same argumentation as in the proof of corollary 8.4.2 ('passing to the slice representation') we find that each $\bar{c}$ is orthogonal as differentiable lift in V with respect to the G-orbits it meets, too.

Suppose $\bar{c}_1 : (a_i, b_i) \supseteq (a, b) \to V\backslash\{0\}$ is a local orthogonal differentiable lift of $c|_{(a_i,b_i)}$ with maximal domain (a, b), where, say, $b < b_i$. Then, there exists a local orthogonal 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. Again our goal is to arrange $\bar{c}_2$ in such a way that we can glue together $\bar{c}_1$ and $\bar{c}_2$ at t_0 differentiably. We may assume without loss that $\bar{c}_1(t_0) = \bar{c}_2(t_0) =: v_0$, by applying a fixed transformation $g \in G$ to $\bar{c}_2$, which preserves the orthogonality of $\bar{c}_2$:

$$0 = \langle \bar{c}_2'(t) | \zeta_X(\bar{c}_2(t)) \rangle = \langle g.\bar{c}_2'(t) | g.\zeta_X(\bar{c}_2(t)) \rangle = \langle g.\bar{c}_2'(t) | \zeta_{\mathrm{Ad}(g).X}(g.\bar{c}_2(t)) \rangle,$$

for all $X \in \mathfrak{g}$, where ζ_X is the fundamental vector field and Ad is the adjoint representation. Since both lifts $\bar{c}_1$ and $\bar{c}_2$ are orthogonal, both derivatives $\bar{c}_1'(t_0)$ and $\bar{c}_2'(t_0)$ lie in the normal subspace N_{v_0}, whence we do not have to care about their $T_{v_0}(G.v_0)$-parts. In the same way as in the proof of theorem 8.3.2 we can find a $g_0 \in G_{v_0}$ such that the derivatives $\bar{c}_1'(t_0)$ and $g_0.\bar{c}_2'(t_0)$ match, and, therefore

$$\bar{c}_{12}(t) := \begin{cases} \bar{c}_1(t) & \text{for} \quad t \le t_0 \\ g_0.\bar{c}_2(t) & \text{for} \quad t \ge t_0 \end{cases}$$

is a local orthogonal differentiable lift of $c|_{(a_i,b_i)}$ defined on a larger interval than $\bar{c}_1$. Note here that, if v_0 is regular, the $\bar{c}_1$ and $\bar{c}_2$ already meet differentiably at v_0, by the uniqueness statement of lemma 7.1.1 or lemma 8.2.1. Consequently, we have shown that each local orthogonal differentiable lift of $c|_{(a_i,b_i)}$ defined on $(a,b) \subseteq (a_i,b_i)$ can be extended to a larger interval whenever $(a,b) \subsetneq (a_i,b_i)$. This proves the claim.

Now let $\bar{c} : (a_i,b_i) \to V\backslash\{0\}$ be the orthogonal differentiable lift of $c|_{(a_i,b_i)}$ constructed above. For $a_i \neq -\infty$, we put $\bar{c}(a_i) := 0$, being the only choice. We are going to extend the lift $\bar{c}$ differentiably to the closure of (a_i,b_i). Consider the expression $\frac{\bar{c}(t)}{t-a_i}$ as $t \searrow a_i$. We want to gain that it is convergent. For t sufficiently close to a_i we have

$$\sigma\left(\frac{\bar{c}(t)}{t-a_i}\right) = c_{(1,a_i)}(t) \to c_{(1,a_i)}(a_i) \qquad \text{as } t \searrow a_i,$$

since $c_{(1,a_i)}(t) := ((t-a_i)^{-2}c_1(t), (t-a_i)^{-d_2}c_2(t), \ldots, (t-a_i)^{-d_n}c_n(t))$ is continuous. Let $\bar{c}_{(1,a_i)}$ be a continuous lift of $c_{(1,a_i)}$ which exists by theorem 8.1.1. It follows that the connected (lemma 8.3.1) set A of all accumulation points of $\left(\frac{\bar{c}(t)}{t-a_i}\right)_{t\searrow a_i}$ lies in the orbit $G.\bar{c}_{(1,a_i)}(a_i)$ through $\bar{c}_{(1,a_i)}(a_i)$. Let us consider the following two cases separately:
If $c_{(1,a_i)}(a_i) = 0$, then $G.\bar{c}_{(1,a_i)}(a_i) = \{0\}$. That means that all accumulation points of $\left(\frac{\bar{c}(t)}{t-a_i}\right)_{t\searrow a_i}$ have to be 0, and so the limit $\lim_{t\searrow a_i} \frac{\bar{c}(t)}{t-a_i}$ exists and equals 0.
If on the other hand $c_{(1,a_i)}(a_i) \neq 0$, then c is normally nonflat at $t = a_i$ which can easily be checked by theorem 7.1.7:

$$(\tilde{\Delta}_1^1 \circ c)(t) = (t-a_i)^2(\tilde{\Delta}_1^1 \circ c_{(1,a_i)})(t) = (t-a_i)^2(\Delta_1^1 \circ \bar{c}_{(1,a_i)})(t),$$

whose right-hand side does only vanish of second order at a_i, since $\Delta_1^1 \circ \bar{c}_{(1,a_i)}$ equals the first coordinate of $c_{(1,a_i)}$ which does not vanish at a_i (if $(\Delta_1^1 \circ \bar{c}_{(1,a_i)})(a_i) = 0$ then $\langle \bar{c}_{(1,a_i)}(a_i) | \bar{c}_{(1,a_i)}(a_i)\rangle = \sigma_1(\bar{c}_{(1,a_i)}(a_i)) = 0$ and so $c_{(1,a_i)}(a_i) = 0$, a contradiction). So we can find an integer r such that the conditions (1) and (2) of theorem 7.1.7 are satisfied. By corollary 8.4.2, there exists a local orthogonal smooth lift $\tilde{c}$ of c near a_i. Moreover, we can find a $t_0 > a_i$ such that both lifts $\bar{c}$ and $\tilde{c}$ are defined near t_0. We can glue together $\bar{c}$ and $\tilde{c}$ differentiably at t_0 by applying a $g_0 \in G$ as above, and the resulting differentiable lift is still orthogonal.
In both cases we have extended the lift $\bar{c}$ differentiably to a_i. Note that for each lift its orthogonality is trivially satisfied at a_i, since each lift has to vanish at a_i and since $N_0 = V$. We can argue the same way for $b_i \neq +\infty$. Thus we have extended $\bar{c}$ differentiably and orthogonally to the closure of (a_i,b_i).
The rest of the proof can be adopted without changes from the proof of theorem 8.3.2: it remains to deal with the points of $c^{-1}(0)$ which we have excluded so far. But at these points each lift is automatically orthogonal, since each lift has to be 0 there. The proof is complete. □

REMARK. In the assumptions of theorem 8.4.3 it is probably sufficient to demand only finite differentiability for the curve c in order to obtain an orthogonal differentiable lift $\bar{c}$. But then it is not clear how to adapt corollary 8.4.2.

8.5. An outlook

For the lifting problem in the special case that the symmetric group S_n is acting on $\mathbb{R}^n$ by permuting the coordinates or, otherwise put (see section 6.1), when smooth parameterizations of the roots of smooth curves of polynomials (of fixed degree n) with only real roots are looked for, we have found strong results in section 5.1:

(1) Any differentiable lift of a C^{2n}-curve (of polynomials) $c : \mathbb{R} \to \mathbb{R}^n/S_n$ is actually C^1.
(2) There always exists a twice differentiable (not better!) lift of c, if it is of class C^{3n}.

Note that here the differentiability conditions of c are not best possible which is shown by the case $n = 2$:

- A C^2-curve (of polynomials) $c : \mathbb{R} \to \mathbb{R}^2/S_2$ allows C^1-lifts.
- There always exists a twice differentiable lift of c, if it is of class C^4.

See proposition 2.1.1 and the remark at the end of section 5.1. Hence there is still room to improve the results (1) and (2).

The proof of (1) and (2) (see theorem 5.1.1) is based on the fact that the roots of a C^n-curve of polynomials $c : \mathbb{R} \to \mathbb{R}^n/S_n$ may be chosen differentiable with locally bounded derivative; this is the content of theorem 3.5.3 and of theorem 4.3.1 due to Bronshtein [**8**] and Wakabayashi [**41**], respectively. To these two theorems are dedicated chapter 3 and chapter 4.

The long-term objective is to transfer the results (1) and (2) to the general situation, primary to prove the existence of a twice differentiable lift in the general setting. And the key is the generalization of Bronshtein's and Wakabayashi's result. This seems to be difficult.

Bibliography

[1] Alekseevsky, Dimitri; Kriegl, Andreas; Losik, Mark; Michor, Peter W., *Choosing roots of polynomials smoothly*, Israel J. Math. **105** (1998), 203-233.

[2] Alekseevsky, Dimitri; Kriegl, Andreas; Losik, Mark; Michor, Peter W., *Lifting smooth curves over invariants for representations of compact Lie groups*, Transformation Groups **5** No. 2 (2000), 103-110.

[3] Apanasov, Boris N., *Discrete groups in space and uniformization problems*, Kluwer Academic Publishers, Dordrecht-Boston-London, 1991.

[4] Baumgärtel, Helmut, *Endlichdimensionale analytische Störungstheorie*, Akademie-Verlag, Berlin, 1972.

[5] Beardon, Alan F., *The geometry of discrete groups*, Springer-Verlag, New York-Heidelberg-Berlin, 1983.

[6] Bierstone, Edward, *Lifting isotopies from orbit spaces*, Topology **14** (1975), 245-252.

[7] Bredon, G.E., *Introduction to compact transformation groups*, Academic Press, New York, 1972.

[8] Bronshtein, M.D., *Smoothness of polynomials depending on parameters*, Sib. Mat. Zh. **20** (1979), 493-501 (Russian); English transl. in Siberian Math. J **20** (1980), 347-352.

[9] Dadok, Jiri, *Polar coordinates induced by actions of compact Lie groups*, Trans. AMS **288** (1985), 125-137.

[10] Dadok, Jiri; Kac, Victor, *Polar representations*, J. Algebra **92** (1985), 504-524.

[11] Dieudonné, J. A., *Foundations of modern analysis, I*, Academic Press, New York - London, 1960.

[12] Dieudonné, J. A., *Sur un theoreme de Glaeser*, J. Anal. Math. **23** (1970), 85-88.

[13] Glaeser, G., *Racine carré d'une fonction différentiable*, Ann. Inst. Fourier (Grenoble) **13,2** (1963), 203-210.

[14] Hörmander, Lars, *Linear Partial Differential Operators*, Springer, Berlin-Göttingen-Heidelberg, 1963.

[15] Hörmander, Lars, *The Analysis of Linear Partial Differential Operators I*, Springer, Berlin-Heidelberg-New York-Tokyo, 1983.

[16] Kempf, G.; Ness, L., *The length of vectors in representation spaces*, Algebraic Geometry, Lect. Notes Math. **732**, 233-243. New York, Springer 1979.

[17] Kriegl, Andreas; Losik, Mark; Michor, Peter W., *Choosing roots of polynomials smoothly, II*, to appear in: Israel J. Math., arXiv:math.CA/0208228.

[18] Kriegl, Andreas; Losik, Mark; Michor, Peter W.; Rainer, Armin, *Lifting mappings over invariants of finite groups*, to appear, arXiv:math.AG/0312030.

[19] Kriegl, Andreas; Losik, Mark; Michor, Peter W.; Rainer, Armin, *Lifting smooth curves over invariants for representations of compact Lie groups, II*, to appear.

[20] Kriegl, Andreas; Michor, Peter W., *The convenient setting of global analysis*, Mathematical Surveys and Monographs, Volume 53, American Mathematical Society, Providence, 1997.

[21] Kolář, I.; Michor, P.W.; Slovák, J., *Natural Operators in Differential Geometry*, Springer Verlag, Berlin, Heidelberg, New York, 1993.

[22] Knapp, Anthony W., *Lie groups beyond an introduction*, Progress in mathematics, v. **140**, Birkhäuser, 1996.

[23] Lehner, Joseph, *Discontinuous groups and automorphic functions*, American Mathematical Society, Providence, Rhode Island, 1964.

[24] Lojasiewicz, S., *Ensembles semi-analytiques*, Inst. Hautes Études Sci.,Bures-sur-Yvette, France (1964).

[25] Luna, D., *Slices étales*, Bull. Soc. Math. Fr. Mém. **33** (1973), 81-105.

[26] Luna, D., *Sur certaines opérations différentiables des groupes de Lie*, Amer. J. Math. **97** (1975), 172-181.

[27] Michor, Peter W., *Isometric actions of Lie groups and invariants*, Lecture notes of a course held in Vienna, 1993, 1997. - http://www.univie.ac.at/~michor/listpubl.html#books

[28] Montgomery, D.; Yang, C.T., *The existence of a slice*, Ann. of Math. **65** (1957), 108-116.
[29] Palais, Richard S., *On the existence of slices for the action of non-compact Lie groups*, Ann. of Math. **73** (1961), 295-323.
[30] Palais, Richard S.; Terng, Chuu-lian, *A general theory of canonical forms*, Trans. AMS **300** (1987), 771-789.
[31] Palais, Richard S.; Terng, Chuu-lian, *Critical point theory and submanifold geometry*, Lecture Notes in Mathematics **1353**, Berlin, Springer-Verlag 1988.
[32] Procesi, Claudio, *Positive symmetric functions*, Adv. Math. **29** (1978), 219-225.
[33] Procesi, Claudio; Schwarz, Gerald, *Inequalities defining orbit spaces*, Invent. math. **81** (1985), 539-554.
[34] Rahman, Q.I.; Schmeisser, G., *Analytic theory of polynomials*, London Mathematical Society Monographs New Series **26**, Oxford Science Publications, 2002.
[35] Rellich, F., *Störungstheorie der Spektralzerlegung, I*, Math. Ann. **113** (1937), 600-619.
[36] Rivlin, T.J, *The Chebyshev polynomials*, Interscience Tracts, Wiley, New York, 1974.
[37] Sartori, G., *A theorem on orbit structures (strata) of compact linear Lie groups*, J. Math. Phys. **24** (1983), 765-768.
[38] Schwarz, Gerald W., *Lifting smooth homotopies of orbit spaces*, Publ. Math. IHES **51** (1980), 37-136.
[39] Schwarz, Gerald W., *Smooth functions invariant under the action of a compact Lie group*, Topology **14** (1975), 63-68.
[40] Terng, Chuu-lian, *Isoparametric submanifolds and their Coxeter groups*, J. Diff. Geom. **21** (1985), 79-107.
[41] Wakabayashi, S., *Remarks on hyperbolic polynomials*, Tsukuba J. Math. **10** (1986), 17-28.
[42] Weyl, Hermann, *The classical groups*, 2nd ed. Princeton: Princeton University Press 1946.

Armin Rainer

CURRICULUM VITAE

I was born on May 21st, 1977 in San Candido/Innichen, Italy. I attended primary school from 1983 to 1988 and secondary school from 1988 to 1991, both in Dobbiaco/Toblach, Italy, my home town. The next five years, 1991 to 1996, I was attending the *Realgymnasium* (high school with scientific orientation) in Brunico/Bruneck, Italy. I started studying mathematics in 1996 at the University of Salzburg, Austria, and I concluded my study with distinction, obtaining the academic degree *Magister der Naturwissenschaften* (Master of Science), in January 2002. During this time I was honored twice with the *Hans Stegbuchner Preis*, once in 2000 for extraordinary achievements as student, once in 2002 for my diploma thesis. Since March 2002 I am working on a doctoral thesis to obtain the degree *Doktor der Naturwissenschaften* (sort of Ph.D) at the University of Vienna, Austria. Starting from July 2002, I am supported by the *Fonds zur Förderung der wissenschaftlichen Forschung* (Austria Science Fund) project number P14195-MAT with employer Peter W. Michor. In August 2003 I attended the two conferences 'Alanfest' and 'Symplectic Geometry and Moment Maps' organized at the Erwin Schrödinger Institute (ESI) in Vienna, Austria. At the '8th Meeting of the Austrian Mathematical Society, Joint Conference in Cooperation with SIMAI and UMI, September 22-26, 2003 in Bolzano, Italy' I held a talk with the title 'Choosing roots of polynomials smoothly and lifting smooth curves over invariants'. In January 2004 I was participant of the '24th Winter School Geometry and Physics' in Srní, Czech Republic.

LIST OF PUBLICATIONS

- *Zerlegungsgleichheit von Kreis und Quadrat*, diploma thesis (in German).
- *Choosing roots of polynomials smoothly and lifting smooth curves over invariants*, doctoral thesis, in progress.
- Kriegl, A., Losik, M., Michor P.W., Rainer A., *Lifting mappings over invariants of finite groups*, to appear, arXiv:math.AG/0312030.
- Kriegl, A., Losik, M., Michor P.W., Rainer A., *Lifting smooth curves over invariants for representations of compact Lie groups II*, to appear.