PERTURBATION THEORY OF POLYNOMIALS AND LINEAR OPERATORS

ADAM PARUSIŃSKI AND ARMIN RAINER

Abstract. This survey revolves around the question how the roots of a monic polynomial (resp. the spectral decomposition of a linear operator), whose coefficients depend in a smooth way on parameters, depend on those parameters. The parameter dependence of the polynomials (resp. operators) ranges from real analytic over C^∞ to differentiable of finite order with often drastically different regularity results for the roots (resp. eigenvalues and eigenvectors). Another interesting point is the difference between the perturbation theory of hyperbolic polynomials (where, by definition, all roots are real) and that of general complex polynomials. The subject, which started with Rellich's work in the 1930s, enjoyed sustained interest through time that intensified in the last two decades, bringing some definitive optimal results. Throughout we try to explain the main proof ideas; Rellich's theorem and Bronshtein's theorem on hyperbolic polynomials are presented with full proofs. The survey is written for readers interested in singularity theory but also for those who intend to apply the results in other fields.

Contents

1. Introduction 2
Notation 5
2. Perturbation theory with real analytic coefficients 6
2.1. Rellich's theorems 6
2.2. A proof of Theorem 2.2 7
2.3. Multiparameter case 10
2.4. Perturbation of normal matrices 11
3. Differentiable roots of hyperbolic polynomials 15
3.1. Bronshtein's theorem 15
3.2. Towards a proof of Bronshtein's theorem 19
3.3. Sufficient conditions for C^p roots 24
3.4. Gårding hyperbolic polynomials 26
3.5. Eigenvalues of Hermitian matrices 29
3.6. Singular values 31
3.7. Eigenvalues of normal matrices 31
4. Regularity of the roots in the general (nonhyperbolic) case 34

Date: October 12, 2023.

2020 *Mathematics Subject Classification.* 26C10, 26B05, 26B30, 26E05, 26E10, 26A46, 30C15, 32B20, 46E35, 47A55.

Key words and phrases. Perturbation theory, regularity of roots of polynomials with smooth coefficients, eigenvalues, eigenvectors, singular values, hyperbolic polynomials, group representations, lifting over invariants, optimal transport.

AR was supported by the FWF-Project P 32905-N.

4.1. The case of radicals 34
4.2. Optimal Sobolev regularity of the roots 35
4.3. Absolute continuity via desingularization. Formulas for the roots 37
4.4. Multiparameter case 38
5. Lifting maps over invariants of group representations 40
5.1. A reformulation of the regularity problem for hyperbolic polynomials 40
5.2. Orthogonal representations of compact Lie groups 40
5.3. The main tools 43
5.4. Examples and applications 45
5.5. A reformulation of the regularity problem for general polynomials 46
5.6. Representations of linearly reductive groups 47
5.7. Some remarks on the proofs 51
6. Applications 52
6.1. Zero sets of smooth functions 52
6.2. Extension to the optimal transport between algebraic hypersurfaces 55
7. Appendix A. Function spaces 57
7.1. Hölder spaces 57
7.2. Lebesgue spaces and weak Lebesgue spaces 58
7.3. Sobolev spaces 59
7.4. Functions of bounded variation 59
8. Appendix B. The space of hyperbolic polynomials 60
Acknowledgements 61
References 61
Index 66

1. Introduction

The question how the spectral decomposition of a linear operator in Hilbert space depends on parameters of the operator is a natural and important one. Its investigation – stimulated without doubt by the development of quantum mechanics – began with Rellich's paper [99] published in 1937 (the first in a series which culminated in Rellich's book [100]). In finite dimensions, the problem reduces to the questions about the regularity of the eigenvalues and eigenvectors of families of matrices and, in consequence, one is led to ask how regular the roots of a monic polynomial depending on parameters can be chosen as functions of these parameters.

In this survey, we will concentrate on the finite dimensional setting of matrices and polynomials; occasionally we will provide references for generalizations to operators in infinite dimensional Hilbert space.

Real analytic perturbations of normal matrices and hyperbolic polynomials. In [99], Rellich was concerned with analytic families of linear operators, in particular, he proved that real analytic curves of Hermitian matrices admit real analytic eigenvalues and a real analytic frame of orthonormal eigenvectors. To this end, he showed that a real analytic curve of monic hyperbolic polynomials admits a real analytic choice of its roots. A monic polynomial $P_a(Z) = Z^d + \sum_{j=1}^d a_j Z^{d-j}$ with real coefficients is called *hyperbolic* if all d roots of P_a (with multiplicities) are real.

Rellich also realized that the analytic dependence of the spectral decomposition breaks down as soon as the Hermitian matrix depends on two or more real parameters. The analytic multiparameter case was taken up and continued by Kurdyka and Păunescu [64]: there exists an analytic modification of the parameter space such that, after this modification, the eigenvalues and the eigenvectors (or the roots in the case of hyperbolic polynomials) admit real analytic parameterizations, locally. Actually, these results hold in the more general setting of normal matrices, as observed by Rainer [94], where the eigenvalues are generally complex valued. We discuss the analytic perturbation theory in Section 2.

Singularity theory methods. Many results of perturbation theory can be proven using singularity theory. Puiseux's theorem is an important ingredient of Rellich's proof and the multiparameter case is proven using its multivariable counterpart, the Abhyankar–Jung theorem. In Section 2 we present a different proof of Rellich's theorem on perturbation of hyperbolic polynomials, based on a nowadays standard singularity theory tool, the splitting, which we also recall briefly. Similar methods are used in the multiparameter case and in the study of the perturbation of normal matrices, see Parusiński and Rond [88], where the splitting is replaced by a version of Hensel's lemma for normal matrices, or in the problem of lifting over invariants of group representations, where it is replaced by the slice theorem (see Theorem 5.9). We describe how this method is used in several places of the survey.

In the multiparameter case, the resolution of singularities is used to make ideals principal generated by normal crossings. In most cases, this is the ideal generated by the discriminant of P_a, or of the square-free reduction $(P_a)_{\mathrm{red}}$, so that one can use then the Abhyankar–Jung theorem. Interestingly, the resolution of singularities is used as well to show that the continuous roots of P_a, with $a = a(t)$ depending smoothly on $t \in \mathbb{R}$, are locally absolutely continuous, see Section 4.3 and Parusiński and Rainer [83].

In the case of C^∞ coefficients (or coefficients of finite differentiability), besides splitting, another important ingredient is Glaeser's inequality and its generalizations. We recall some of its manifestations in Section 3. The higher order Glaeser inequalities appear in Section 4.

C^∞ perturbations of normal matrices and hyperbolic polynomials. Another field of application which greatly stimulated the research on the regularity problem for roots of polynomials is the theory of PDEs – most notably the Cauchy problem for hyperbolic PDOs. With this goal in mind, Bronshtein [21] proved that any continuous root of a monic hyperbolic polynomial P_a of degree d with coefficients $C^{d-1,1}$ functions $a_j : U \to \mathbb{R}$, where U is an open subset of $\mathbb{R}^n$, is locally Lipschitz. (Glaeser's inequality gives Bronshtein's result in the simplest nontrivial case.) Note that ordering the roots P_a increasingly yields a continuous system of the roots. This result is sharp. If $n = 1$ and the coefficients are C^d (resp. C^{2d}), then there is a C^1 (resp. twice differentiable) choice of the roots, but there are examples of nonnegative C^∞ functions f such that $Z^2 = f$ has no $C^{1,\alpha}$ solution for any $\alpha > 0$. We give a full proof of Bronshtein's theorem with uniform bounds for the Lipschitz constant in Section 3.2 (it is based on Parusiński and Rainer [81]).

The step from analytic to C^∞ coefficients entails a big drop of regularity for the roots, because it allows for oscillation. The same is true for the eigenvalues of Hermitian and, more generally, normal matrices; for the eigenvectors it is even

worse: there are C^∞ curves of symmetric 2×2 matrices that do not admit a continuous system of eigenvectors. On the other hand, the conditions that guarantee locally Lipschitz or C^1 eigenvalues of Hermitian (even normal) matrices are much weaker than for hyperbolic polynomials: for instance, a $C^{0,1}$ (resp. C^1) curve of Hermitian (or, more generally, normal) matrices admits a system of $C^{0,1}$ (resp. C^1) eigenvalues. We explore several explanations for this phenomenon: one of them is that the determinant is a Gårding hyperbolic polynomial with respect to the identity matrix on the real vector space of $d \times d$ Hermitian matrices (see Section 3.4 and Section 3.5). This also gives that $C^{1,1}$ curves of Hermitian matrices admit eigenvalues that are locally of Sobolev class $W^{2,1}$ (which is optimal in view of the Sobolev inequality).

There are different approaches (discussed in Section 3.3) to ascertain better C^p regularity of the roots and the eigenvalues by imposing stronger assumptions. For instance, given any integer $p \geq 2$, forcing $f \geq 0$ defined on $\mathbb{R}$ and sufficiently many of its derivatives to vanish on all local minima of f guarantees that the equation $Z^2 = f$ has a C^p solution. Or, one can define a regularity class $\mathcal{F}^\beta$, $\beta > 0$, of nonnegative C^β functions on $\mathbb{R}$, incorporating flatness of f near its zeros in a certain sense, such that $f^\alpha \in \mathcal{F}^{\alpha\beta}$ for all $f \in \mathcal{F}^\beta$ and $\alpha \in (0, 1]$. In the case of a curve of monic hyperbolic polynomials or normal matrices, sufficient conditions for the existence of C^p roots or eigenvalues can be given in terms of the differentiability of the coefficients and the finite order of contact of the roots or eigenvalues. If the coefficients are C^∞ and definable in any (not necessarily polynomially bounded) o-minimal expansion of the real field, then the roots and eigenvalues can be chosen C^∞ and definable; this reaffirms that oscillation is to blame for the loss of regularity.

Perturbations of polynomials. General (nonhyperbolic) case. Other problems from PDEs (cf. [106]) and geometric analysis demand to abandon the hyperbolicity assumption: how regular can the roots of a general (nonhyperbolic) monic polynomial P_a with smooth coefficients $a_j : U \to \mathbb{C}$ be? This problem was solved only recently by Ghisi and Gobbino [37] in the radical case and by Parusiński and Rainer [83, 84, 85] in the general case. We present in Section 4 two approaches to this problem: one is based on formulas for the roots of the universal monic polynomial of degree d, obtained by resolution of singularities, the other one is elementary and yields the optimal Sobolev regularity result. It states that the roots of a curve of monic complex polynomials P_a of degree d with coefficients of class $C^{d-1,1}(\bar{I})$, where $I \subseteq \mathbb{R}$ is a bounded open interval, admit an absolutely continuous parameterization whose derivative is in $L^p(I)$ for each $1 \leq p < d/(d-1)$. Actually, each continuous root has this property, in particular, it belongs to the Sobolev space $W^{1,p}(I)$ for each $1 \leq p < d/(d-1)$. Moreover, there are uniform bounds for the Sobolev norm of the roots in terms of the $C^{d-1,1}$ norm of the coefficients. This result is optimal.

A multiparameter version for continuous roots follows by standard arguments, provided there exists a continuous root (perhaps on some subset). But in general there are obstructions for continuous roots, due to monodromy. It turns out that the possibly discontinuous roots still can be represented by functions of bounded variation, cf. Section 4.4.1. The proof uses the formulas for the roots of the universal monic polynomial alluded to above.

Lifting from the orbit space. The problem of finding roots with optimal regularity of monic (hyperbolic and nonhyperbolic) polynomials has a representation-theoretic interpretation: it is equivalent to a lifting problem over the orbit map σ of the standard representation of the symmetric group S_d on $\mathbb{R}^d$ or $\mathbb{C}^d$, respectively, where the components of σ are the elementary symmetric functions. This point of view can be generalized considerably: we will discuss in Section 5 the lifting problem for real orthogonal representations of compact Lie groups and for complex rational representations of linearly reductive groups (the representation spaces will always be finite dimensional). In this general setting, the lifting occurs over a map σ consisting of basic invariants, that is a finite collection of generators of the algebra of invariant polynomials. It gives another heuristic explanation why the eigenvalues of symmetric matrices are better behaved than the roots of monic hyperbolic polynomials: only a "partial lifting" is necessary (see Section 5.4.1).

Applications. In the last section, Section 6, we give two applications of the results for general (nonhyperbolic) polynomials. The first concerns the zero set of C^∞ functions f. If f vanishes to some finite order γ at a point $x_0 \in \mathbb{R}^n$, then locally near x_0 the zero set of f is given by the real roots of a monic polynomial P_a of degree γ, by Malgrange's preparation theorem. We present several consequences of the optimal Sobolev regularity of the roots of P_a, most of them observed by Beck, Becker-Kahn, and Hanin [10]. Often one knows in advance the vanishing order of a function so that this setup applies, e.g., for solutions of second order elliptic equations, Laplace eigenfunctions or finite linear combinations of such.

The second application is taken from the recent paper [5] by Antonini, Cavalletti, and Lerario. This paper brings forward a reinterpretation and an extension of the regularity problem for the roots of polynomials to the study of the Wasserstein distance on the space of d-degree hypersurfaces in $\mathbb{CP}^n$. This space is embedded in the space of measures on the projective space and thus the optimal transport problem between hypersurfaces is studied. An inner Wasserstein distance is obtained which turns the projective space of homogeneous polynomials in a complete geodesic space.

In the appendix, we collect definitions and basics on function spaces used in the survey (Section 7) and we present some topological and geometric properties of the space of monic hyperbolic polynomials of a fixed degree (Section 8). There are several open problems that we included all over the article.

Notation

We use $\mathbb{N} := \{0, 1, 2, \ldots\}$ and $\mathbb{N}_{\geq m} := \{n \in \mathbb{N} : n \geq m\}$ for $m \in \mathbb{N}$. Similarly, $\mathbb{R}_{>a} := \{x \in \mathbb{R} : x > a\}$ etc.; for instance $\mathbb{R}_{\geq 0} = [0, \infty)$ and $\mathbb{R}_{>0} = (0, \infty)$.

For a subset S of a topological space, S°, $\overline{S}$, and ∂S denote the interior, the closure, and the boundary of S, respectively. We consider $\mathbb{R}^n$ with its Euclidean structure $|x| = \sqrt{x_1^2 + \cdots + x_n^2}$, if not stated otherwise. The open ball in $\mathbb{R}^n$ with center x and radius $r > 0$ is $B(x, r) := \{y \in \mathbb{R}^n : |x - y| < r\}$; if $n = 1$, we use $I(x, r) := B(x, r)$. The unit sphere in $\mathbb{R}^n$ is denoted by $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n : |x| = 1\}$. If $U \subseteq \mathbb{R}^n$ is an open subset of $\mathbb{R}^n$, then we write $V \Subset U$ to denote a relatively compact open subset V of U.

The Lebesgue measure in $\mathbb{R}^n$ is denoted by $\mathcal{L}^n$; we also use $|E| := \mathcal{L}^n(E)$ for a Lebesgue measurable set $E \subseteq \mathbb{R}^n$. The k-dimensional Hausdorff measure is $\mathcal{H}^k$

and $\mathcal{H}^k \llcorner E$ denotes the restriction of $\mathcal{H}^k$ to a subset $E \subseteq \mathbb{R}^n$, i.e., $(\mathcal{H}^k \llcorner E)(F) := \mathcal{H}^k(E \cap F)$.

For a function f, we write Z_f for its zero set and Γ_f for its graph. If f and g are real valued functions, $f \lesssim g$ means that $f \le C\, g$ for some constant $C > 0$. Generic constants are denoted by C or $C(n, d, \dots)$ to indicate that it depends on $n, d, \dots$; their value may differ from line to line.

We use standard multiindex notation. We write df for the total derivative and $d_v f(x) := \partial_t|_{t=0} f(x+tv)$ for the directional derivative. In particular, $\partial_j f = \frac{\partial}{\partial x_j} f = d_{e_j} f$, where e_j is the j-th standard unit vector in $\mathbb{R}^n$. We write $\nabla f = (\partial_1 f, \dots, \partial_n f)$ for the gradient. Higher order partial derivatives are denoted by $f^{(\alpha)} = \partial^\alpha f = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} f$, for $\alpha \in \mathbb{N}^n$.

A monic polynomial of degree d is $P_a(Z) = Z^d + \sum_{j=1}^d a_j Z^{d-j}$, where the subscript indicates the vector of coefficients $a = (a_1, \dots, a_d)$. The coefficients may be real or complex numbers, formal power series, or functions; in the latter cases we often write $P_a(X)(Z)$, $P_a(t)(Z)$, or $P_a(x)(Z)$ to indicate the variables on which the coefficients depend. If the first coefficient a_1 is zero, then the polynomial is said to be in Tschirnhausen form and in that case we consistently denote its coefficients by $\tilde a_j$, i.e., $P_{\tilde a}(Z) = Z^d + \sum_{j=2}^d \tilde a_j Z^{d-j}$.

By a parameterization or a system of the roots of a family of polynomials $P_a(x)$, where x ranges over a subset U of $\mathbb{R}^n$, we mean a collection of d functions $\lambda_1, \dots, \lambda_d$ such that $P_a(x)(Z) = \prod_{i=1}^d (Z - \lambda_i(x))$ for all $x \in U$.

The space of monic hyperbolic polynomials of degree d is denoted by $\mathrm{Hyp}(d)$, its subspace of polynomials in Tschirnhausen form by $\mathrm{Hyp}_T(d)$, and $\mathrm{Hyp}_T^0(d) := \{P_{\tilde a} \in \mathrm{Hyp}_T(d) : \tilde a_2 = -1\}$.

For a polynomial $P(Z)$, we denote by Δ_P its discriminant. For an analytic function or a formal power series f, we denote by f_{red} its square-free reduction. For a square matrix A, we denote by χ_A its characteristic polynomial and by Δ_A the discriminant of $(\chi_A)_{\mathrm{red}}$.

We write $\mathrm{Mat}_{m,n}(R)$ for the set of $m \times n$ matrices over the ring R and set $\mathrm{Mat}_n(R) := \mathrm{Mat}_{n,n}(R)$; for instance, $\mathrm{Mat}_{m,n}(\mathbb{K}[[X]])$ is the set of $m \times n$ matrices with entries formal power series in X with coefficients in $\mathbb{K}$. For $R = \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ we also use $\mathbb{K}^{m\times n} := \mathrm{Mat}_{m,n}(\mathbb{K})$. For $A \in \mathrm{Mat}_{m,n}(\mathbb{C})$ let $A^* = \overline{A}^t$ be its conjugate transpose. The identity matrix is denoted by $\mathbb{I}$; its size will always be clear from the context. We write $\mathrm{U}(n) = \mathrm{U}_n(\mathbb{C})$ for the group of unitary matrices and $\mathrm{O}(n) = \mathrm{O}_n(\mathbb{R})$ for the group of orthogonal matrices. More abstractly, $\mathrm{O}(V)$ is the group of orthogonal endomorphisms of a finite dimensional Euclidean vector space V. $\mathrm{GL}_n(\mathbb{C})$ or $\mathrm{GL}_n(\mathbb{R})$ denotes the general linear group. The set of diagonal $n \times n$ matrices is denoted by $\mathrm{Diag}(n)$, the set of real symmetric $n \times n$ matrices by $\mathrm{Sym}(n)$ and the set of complex Hermitian $n \times n$ matrices by $\mathrm{Herm}(n)$.

The symmetric group of permutations of n elements is S_n. It acts on $\mathbb{K}^n$, where $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, by permuting the coordinates and thus induces an equivalence relation. The equivalence class of $x = (x_1, \dots, x_n) \in \mathbb{K}^n$ can be identified with the unordered n-tuple $[x] = [x_1, \dots, x_n]$ (which may be represented by $\sum_{j=1}^n \delta_{x_j}$, where δ_{x_j} is the Dirac measure).

2. Perturbation theory with real analytic coefficients

2.1. **Rellich's theorems.** In 1937 F. Rellich proved the following result.

Theorem 2.1 (Rellich [99, Satz I], [100, Theorem 1]). *Let $A(t)$ be a family of $d \times d$ Hermitian matrices whose coefficients are real analytic in $t \in I$, where I is an open neighborhood of 0 in $\mathbb{R}$. Then, maybe in a smaller neighborhood of 0, the eigenvalues and the eigenvectors of $A(t)$ can be chosen real analytic in t.*
More precisely, there exist real valued functions $\lambda_i(t)$ and complex valued d-vectors $v_i(t)$, $i = 1, \dots, d$, depending analytically on t, such that for t in an open neighborhood I' of 0 in $\mathbb{R}$:

(1) *$v_1(t), \dots, v_d(t)$ is an orthonormal frame in $\mathbb{C}^d$;*
(2) *$A(t)v_i(t) = \lambda_i(t)v_i(t)$.*

Note that the statement of this theorem is local. It was originally stated and proven in terms of convergent power series but clearly, by unique analytic continuation, it implies a global statement on an arbitrary open interval I and with $I' = I$. For proving the above theorem, Rellich shows a result that we rephrase as follows.

Theorem 2.2 (Rellich [99]). *Let $I \subseteq \mathbb{R}$ be an open neighborhood of 0 in $\mathbb{R}$ and let*

$$P_a(t)(Z) = Z^d + \sum_{j=1}^{d} a_j(t) Z^{d-j}, \quad t \in I,$$

be a monic hyperbolic polynomial with real analytic coefficients a_j. Then there exist real analytic functions λ_i, $1 \le i \le d$, defined maybe in a smaller neighborhood I' of 0 in $\mathbb{R}$, which parameterize the roots of P_a, i.e.,

$$P_a(t)(Z) = \prod_{i=1}^{d} (Z - \lambda_i(t)), \quad t \in I'.$$

Actually, Rellich states this result for complex analytic functions that have real values for real t, see Rellich [99, Hilfssatz II]. Again, for real t, it is easy to see that it holds globally, that is for an arbitrary I and with $I' = I$.

Rellich's proof of Theorem 2.2 is based on Puiseux's theorem. Its idea goes as follows. Let $\lambda_i(t)$ be the complex roots of $P_a(t)$. By Puiseux's theorem, there is $l \in \mathbb{N}$ such that $\eta_i(s) = \lambda_i(s^l)$ are all complex analytic, but also, by the assumption, every $\eta_i(s)$ is real whenever s^l is real. But this means that η_i, as a power series, depends only on s^l. Indeed, suppose this is not true and let $c_k s^k$ be the lowest degree term in the expansion of η_i such that l does not divide k. Then η_i can not be real for all s such that s^l is real. (Rellich gives a slightly different argument still based on Puiseux's theorem in [100, pages 30-31].)

Then to show Theorem 2.1 Rellich, for each analytic eigenvalue $\lambda(t)$ of $A(t)$, solves the equation $A(t)v = \lambda(t)v$ using formulas involving the minors of the matrix $A(t)$. We refer the reader to Rellich's paper for this elementary argument.

Remark 2.3. It was observed by Rainer [94] that Theorem 2.1 holds as well for normal matrices. This shows, surprisingly, that it is not the hyperbolicity of the characteristic polynomial of $A(t)$ that is important but the fact that we can diagonalize the matrices in orthonormal bases, but see also Example 2.16. This was made even more transparent in the proof for normal matrices given in [88], where analytic diagonalizibility is showed directly without proving first the analyticity of the eigenvalues. See also Section 2.4.

2.2. **A proof of Theorem 2.2.** We present here a different proof based on the splitting principle.

2.2.1. *Tschirnhausen form.* Every monic real polynomial

$$P_a(Z) = Z^d + \sum_{j=1}^{d} a_j Z^{d-j}$$

of degree d can be identified with the point $a = (a_1, \dots, a_d) \in \mathbb{R}^d$. We say that P_a is in *Tschirnhausen form* if $a_1 = 0$. Every P_a can be put in Tschirnhausen form by the substitution

$$P_{\tilde{a}}(Z) = P_a(Z - \tfrac{a_1}{d}) = Z^d + \sum_{j=2}^{d} \tilde{a}_j Z^{d-j}.$$

We call it *Tschirnhausen transformation*. Note that

$$\tilde{a}_j = \sum_{i=0}^{j} C_i a_i a_1^{j-i}, \quad 2 \le j \le d, \tag{1}$$

where the C_i are universal constants and $a_0 = 1$. For a polynomial $P_{\tilde{a}}$ in Tschirnhausen form we have

$$-2\tilde{a}_2 = \lambda_1^2 + \dots + \lambda_d^2,$$

where $\lambda_1, \dots, \lambda_d$ are the roots of $P_{\tilde{a}}$ (with multiplicities), consequently, for a hyperbolic polynomial,

$$\tilde{a}_2 \le 0.$$

Recall that the coefficients (up to their sign) are the elementary symmetric polynomials in the roots, by Vieta's formulas.

Lemma 2.4. *The coefficients of a hyperbolic polynomial $P_{\tilde{a}}$ in Tschirnhausen form satisfy*

$$|\tilde{a}_j|^{1/j} \le \sqrt{2}\, |\tilde{a}_2|^{1/2}, \quad j = 1, \dots, d. \tag{2}$$

Proof. Let $s_k := \lambda_1^k + \dots + \lambda_d^k$, $k \ge 1$, be the Newton polynomials in $\lambda_1, \dots, \lambda_d$. We claim that

$$|s_k|^{1/k} \le |s_2|^{1/2}, \quad k = 2, \dots, d. \tag{3}$$

Setting $\lambda = (\lambda_1, \dots, \lambda_d)$, we have, for $k \ge 2$,

$$|s_k|^{1/k} \le \|\lambda\|_k \le \|\lambda\|_2 = |s_2|^{1/2},$$

by a well-known relation between the p-norms. The Newton identities between the Newton polynomials and the elementary symmetric polynomials σ_k,

$$j\sigma_j = \sum_{i=1}^{j} (-1)^{i-1} \sigma_{j-i} s_i, \quad d \ge j \ge 1,$$

imply

$$|\tilde{a}_j| \le \frac{1}{j} \sum_{i=2}^{j} |\tilde{a}_{j-i}||s_i| \le \frac{1}{j} \sum_{i=2}^{j} |\tilde{a}_{j-i}||s_2|^{i/2}.$$

By (3), it is now easy to conclude the statement using induction. □

2.2.2. *Splitting.* The following well-known lemma can be found for instance in [14].

Lemma 2.5. *Let $P_a = P_b P_c$, where P_b and P_c are monic real polynomials without common (complex) root. Then we have $P = P_{b(P)} P_{c(P)}$ for analytic mappings $P \mapsto b(P) \in \mathbb{R}^{\deg P_b}$ and $P \mapsto c(P) \in \mathbb{R}^{\deg P_c}$, defined for P near P_a in $\mathbb{R}^{\deg P_a}$, with the given initial values.*

Proof. The product $P_a = P_a P_c$ defines on the coefficients a polynomial map φ such that $a = \varphi(b, c)$. Its Jacobian determinant $\det \mathrm{Jac}_\varphi(b, c)$ equals the resultant of P_b and P_c which is nonzero, by assumption. Thus φ can be inverted locally, by the inverse function theorem. □

Let $\mathrm{Hyp}_T(d)$ be the space of monic hyperbolic polynomials of degree d in Tschirnhausen form and let $\mathrm{Hyp}_T^0(d)$ be the subspace of polynomials $P_{\tilde{a}}$ with $\tilde{a}_2 = -1$. By Lemma 2.4, $\mathrm{Hyp}_T^0(d)$ is compact. Every polynomial in $\mathrm{Hyp}_T^0(d)$ can be written as a product of two monic real polynomials of positive degree without common root (indeed, the sum of the roots vanishes while the sum of their squares is 2).

Let us now prove Theorem 2.2. The case $\tilde{a}_2 = 0$ is trivial, since then all $\tilde{a}_j = 0$. Thus let $P_{\tilde{a}} \in \mathrm{Hyp}_T(d)$ be such that $\tilde{a}_2 \neq 0$. Since $-\tilde{a}_2(t)$ is positive, it is of the form $-\tilde{a}_2(t) = t^{2k} u(t)$, $u(0) > 0$. Choose one of the two analytic roots of $-\tilde{a}_2(t)$, $\pm t^k u^{\frac{1}{2}}(t)$, and denote it by $\theta(t)$. Then the polynomial

$$Q_{\underline{a}}(Z) := \theta^{-d} P_{\tilde{a}}(\theta Z) = Z^d - Z^{d-2} + \sum_{j=3}^{d} \theta^{-j} \tilde{a}_j Z^{d-j}$$

belongs to $\mathrm{Hyp}_T^0(d)$. By Lemma 2.5, we have

$$Q_{\underline{a}} = Q_{\underline{b}} Q_{\underline{c}},$$

on some open ball $B(P_{\tilde{a}}, r) \subseteq \mathbb{R}^d$ such that $\deg Q_{\underline{b}} < d$, $\deg Q_{\underline{c}} < d$, and

$$\underline{b}_i = \psi_i(\theta^{-3}\tilde{a}_3, \dots, \theta^{-d}\tilde{a}_d), \quad i = 1, \dots, \deg Q_{\underline{b}},$$

where ψ_i are real analytic functions; likewise for $\underline{c}_i$. Note that, by Lemma 2.4, $\theta^{-i}\tilde{a}_i$ are real analytic and hence so are b_i and c_i. If $Q_{\underline{a}}$ is hyperbolic, then also $Q_{\underline{b}}$ and $Q_{\underline{c}}$ are hyperbolic. This induces a splitting

$$P_{\tilde{a}} = P_b P_c, \quad \text{on } B(P_{\tilde{a}}, r), \tag{4}$$

where

$$b_i = \theta^i \psi_i(\theta^{-3}\tilde{a}_3, \dots, \theta^{-d}\tilde{a}_d), \quad i = 1, \dots, \deg P_b,$$

are real analytic functions of t and likewise for c_i. To conclude we proceed by induction on $\deg P_a$. □

Remark 2.6. At first the above proof looks more complicated than Rellich's original argument. This is because Rellich uses Puiseux's theorem, that can be proven itself using the splitting. But the splitting allows to work with coefficients a_i that are not necessarily real analytic, but C^p or C^∞ for instance, and that cannot be easily complexified. It can be also be used in the multiparameter case.

It is natural to ask what happens if the entries of a Hermitian matrix are C^∞ functions of $t \in \mathbb{R}$. Then, in general, there is a drastic drop of regularity for the eigenvalues (this will be discussed in detail in Section 3.5 and Section 3.7), while the eigenvectors may not even admit a continuous choice, as seen in the following example, see also Example 3.48.

Example 2.7 ([99, §2.2])**.** The C^∞ curve of symmetric matrices

$$A(t) := e^{-1/t^2}\begin{pmatrix}\cos t^{-1} & \sin t^{-1}\\ \sin t^{-1} & -\cos t^{-1}\end{pmatrix}, \quad t\neq 0,$$

and $A(0) := 0$ has the C^∞ eigenvalues $\pm e^{-1/t^2}$, but the normalized eigenvectors cannot be continuous at $t = 0$.

Under the additional assumption that no two continuous eigenvalues meet to infinite order of flatness, Rellich's theorem 2.1 still holds in the C^∞ category; see [2, Theorem 7.6] and [94] for normal matrices.

2.3. **Multiparameter case.** Both Rellich's theorems fail in the multiparameter case.

Example 2.8 (Rellich [99, §2, 1])**.** The eigenvalues of the symmetric matrix

$$A(x_1,x_2) = \begin{pmatrix}1+2x_1 & x_1+x_2\\ x_1+x_2 & 1+2x_2\end{pmatrix},$$

$\lambda_\pm = 1 + x_1 + x_2 \pm \sqrt{2(x_1^2 + x_2^2)}$, are not analytic in any neighborhood of $(0, 0)$.

In [64] K. Kurdyka and L. Păunescu proposed multiparameter versions of Rellich's theorems. The following is a multiparameter generalization of Theorem 2.2.

Theorem 2.9 ([64, Theorems 4.1 and 5.8])**.** *Let $U \subseteq \mathbb{R}^n$ be open and*

$$P_a(x)(Z) = Z^d + \sum_{j=1}^d a_j(x)Z^{d-j}, \quad x\in U,$$

be a monic hyperbolic polynomial with real analytic coefficients a_j. Then there exist locally Lipschitz functions $\lambda_i : U \to \mathbb{R}$, $i = 1, \dots, d$, which parameterize the roots of P_a.

Moreover, there exists a modification $\sigma : W \to U$ (by a locally finite composition of blowings-up with global smooth centers) such that the roots of $P_a(\sigma(w))(Z)$ can be parameterized, locally on W, by real analytic functions.

Since all roots of $P_a(x)$ are real, we can order them

$$\lambda_1(x) \le \lambda_2(x) \le \dots \le \lambda_d(x)$$

and the first claim of the above theorem implies that all such λ_i are locally Lipschitz.

Example 2.10 ([64, Example 5.9])**.** Consider $P_a(x)(Z) = Z^2 - (x_1^2 + x_2^2)$. The zero set of P_a is the double cone. Then both the upper and the lower nappes are graphs of Lipschitz functions. But note that if we restrict them to a line through the origin these restrictions are not real analytic. We have to interchange them while passing through the origin to get analytic roots (that exist by Rellich's theorem).

If we blow up the origin in the real plane, then the roots become locally, but not globally, real analytic. Indeed, in the chart $x_1 = w_1, x_2 = w_1w_2$ the roots are $\pm w_1\sqrt{1 + w_2^2}$. Going along the exceptional divisor, the center circle of the Möbius band, changes the sign of the root, i.e., interchanges them.

To show the second claim of Theorem 2.9 Kurdyka and Păunescu use resolution of singularities, splitting, and the Abhyankar–Jung theorem.

2.3.1. *Abhyankar–Jung theorem.* Let $\mathbb{K}$ be a field and let

$$P(Z) = Z^d + a_1(X)Z^{d-1} + \cdots + a_d(X) \in \mathbb{K}[[X]][Z] \tag{5}$$

be a monic polynomial whose coefficients are formal power series in $X = (X_1, \ldots, X_n)$. Such a polynomial P is called *quasi-ordinary* if the discriminant $\Delta_P(X)$ equals $X_1^{\alpha_1} \cdots X_n^{\alpha_n} U(X)$, with $\alpha_i \in \mathbb{N}$ and $U(0) \neq 0$.

Theorem 2.11 (Abhyankar–Jung theorem)**.** *Let $\mathbb{K}$ be an algebraically closed field of characteristic zero and let $P \in \mathbb{K}[[X]][Z]$ be a quasi-ordinary polynomial such that the discriminant of P is of the form $\Delta_P(X) = X_1^{\alpha_1} \cdots X_r^{\alpha_r} U(X)$, where $U(0) \neq 0$, and $r \leq n$. Then there is $q \in \mathbb{N} \setminus \{0\}$ such that $P(Z)$ has its roots in $\mathbb{K}[[X_1^{\frac{1}{q}}, ..., X_r^{\frac{1}{q}}, X_{r+1}, ..., X_n]]$.*

The Abhyankar–Jung theorem can be understood as a multiparameter version of Puiseux's theorem. It has first been proven by Jung in 1908 for $n = 2$ and $\mathbb{K} = \mathbb{C}$, cf. [51]. The first complete proof of Theorem 2.11 appeared in [1].

The Abhyankar–Jung theorem still holds if P has multiple factors (so its discriminant is identically equal to zero) but the discriminant of the square-free reduction P_{red} of P is of the form $X_1^{\alpha_1} \cdots X_r^{\alpha_r} U(X)$.

The Abhyankar–Jung theorem can be proven using splitting, as it is done in [87]. But in order to apply the splitting one needs the following result that is much more difficult to prove, see [87, Theorem 1.1 and Remark 1.4].

Theorem 2.12. *Let $\mathbb{K}$ be a (not necessarily algebraically closed) field of characteristic zero and let $P \in \mathbb{K}[[X]][Z]$ be a quasi-ordinary polynomial in Tschirnhausen form (i.e. $a_1 = 0$). Then the ideal $(a_i^{d!/i}(X))_{i=2,\ldots,d}$ is generated by one of $a_i^{d!/i}(X)$ and this generator equals a monomial in $X_1, \ldots, X_n$ times a unit.*

2.3.2. *Proof of Theorem 2.9.* We sketch a proof of this theorem.

The first claim of the theorem follows from Bronshtein's theorem, see Theorem 3.2 below. The proof given in [64] is different and based on real analytic geometry arguments.

To show the second claim one first applies Hironaka's resolution of singularites to make the discriminant $\Delta_{P_{\text{red}}}$ of P_{red} normal crossings. For this it suffices, by [49, Main Theorem II"(N), page 158], to apply a modification given by the composition of a locally finite number of global blowings-up with smooth centers. Therefore, one may suppose that P_{red} is quasi-ordinary. To conclude Kurdyka and Păunescu use the Abhyankar–Jung theorem applied to a hyperbolic quasi-ordinary polynomial, see [64, Proposition 5.4] for details. Instead, one may use here directly Theorem 2.12 and splitting. Firstly, after applying the Tschirnhausen transformation we may suppose $a_1 = 0$. Secondly, by Lemma 2.4, it is $a_2^{d!/2}$ that generates the ideal $(a_i^{d!/i})_{i=2,\ldots,d}$. Denote by $\theta(x)$ one of the two analytic square roots of $-a_2(x)$, that is, if $-a_2(x) = u(x) \prod x_k^{2\alpha_k}$, then $\theta(x) = \pm u^{\frac{1}{2}}(x) \prod x_k^{\alpha_k}$. Then, one concludes exactly as in the proof of Rellich's theorem by splitting the polynomial P and proceeding by induction on its degree. □

2.4. **Perturbation of normal matrices.** Kurdyka and Păunescu showed multiparameter counterparts of Theorem 2.1 for Hermitian matrices.

Theorem 2.13 ([64, Theorem 6.2 and Remark 6.3])**.** *Let $U \subseteq \mathbb{R}^n$ be open and let $A(x)$, $x \in U$, be an analytic family of Hermitian matrices. Then, there exists*

a locally finite composition of blowings-up with smooth global centers $\sigma : W \to U$, such that locally on W the corresponding family $A\circ\sigma$ admits a simultaneous analytic diagonalization.

They also gave a version for real antisymmetric matrices, [64, Theorem 7.2]. This was extended to normal matrices by Rainer [94] by showing that the eigenvalues and eigenspaces of normal matrices depend analytically on the parameters after a modification of the parameter space. Note that these results generalize the one parameter case because there are no nontrivial modifications of one dimensional nonsingular spaces.

Recall that for a square matrix A we denote by Δ_A the discriminant of the reduced (i.e., square-free) form $(\chi_A)_{\mathrm{red}}$ of its characteristic polynomial χ_A. The normal crossings assumption for Δ_A is sufficient for the local simultaneous reduction of families of normal matrices as it is showed by Parusiński and Rond in [88]. The following three algebraic theorems, stated in terms of formal power series $\mathbb{K}[[X]]$, $\mathbb{K} = \mathbb{R}, \mathbb{C}$, were proven in [88]. They hold as well for several subrings of $\mathbb{K}[[X]]$ including the algebraic power series $\mathbb{K}\langle X\rangle$ and the convergent power series $\mathbb{K}\{X\}$, see Remark 2.20.

Theorem 2.14 ([88, Theorem 2.5])**.** *Let $A(X) \in \mathrm{Mat}_d(\mathbb{C}[[X]])$ be normal and suppose that $\Delta_A(X) = X_1^{\alpha_1}\cdots X_n^{\alpha_n}u(X)$ with $u(0) \neq 0$. Then there is a unitary matrix $U(X) \in \mathrm{U}_d(\mathbb{C}[[X]])$ such that*

$$U(X)^{-1}A(X)U(X) = D(X),$$

where $D(X)$ is a diagonal matrix with entries in $\mathbb{C}[[X]]$.

Theorem 2.15 ([88, Theorem 2.9])**.** *Let $A(X) \in \mathrm{Mat}_d(\mathbb{R}[[X]])$ be normal and suppose that $\Delta_A(X) = X_1^{\alpha_1}\cdots X_n^{\alpha_n}u(X)$ with $u(0) \neq 0$. Then there exists an orthogonal matrix $O(X) \in \mathrm{O}_d(\mathbb{R}[[X]])$ such that $O(X)^{-1}A(X)O(X)$ is block diagonal with the blocks of size 1 and 2. The blocks of size 2 are 2×2 matrices of the form*

$$\begin{pmatrix} a(X) & b(X) \\ -b(X) & a(X) \end{pmatrix} \tag{6}$$

for some $a(X)$, $b(X) \in \mathbb{R}[[X]]$.

Note that it follows that if $A(X)$ is symmetric, resp. antisymmetric, then $O(X)^{-1}A(X)O(X)$ is symmetric (i.e., diagonal), resp. antisymmetric.

Example 2.16 (Rainer [94, Example 8.3])**.** The eigenvalues of the one parameter diagonalizable, but not normal, matrix

$$A(t) = \begin{pmatrix} t & 0 & 0 \\ 0 & 0 & t^2 \\ 0 & t & 0 \end{pmatrix}, \quad t \in \mathbb{R},$$

are $t, \pm t^{3/2}$ for $t \geq 0$ and $t, \pm i|t|^{3/2}$ for $t < 0$.

The above example shows that the eigenvalues of a real analytic family of diagonalizable matrices are not necessarily analytic (even for one parameter families). As the following theorem shows, for arbitrary analytic families of matrices, not necessarily square, the counterparts of the above two theorems hold for the singular value decomposition.

Let $A \in \mathrm{Mat}_{m,n}(\mathbb{C})$. Then $A = VDU^{-1}$ for some unitary matrices $V \in \mathrm{U}_m(\mathbb{C})$, $U \in \mathrm{U}_n(\mathbb{C})$, and a (rectangular) diagonal matrix D with real nonnegative coefficients. The diagonal elements of D are the nonnegative square roots of the eigenvalues of A^*A or, equivalently, of AA^*. They are called *singular values* of A. If A is real, then V and U can be chosen orthogonal.

Theorem 2.17 ([88, Theorems 3.1 and 3.3])**.** *Let $A = A(X) \in \mathrm{Mat}_{m,d}(\mathbb{C}[[X]])$, $d \leq k$, and suppose that $\Delta_{A^*A}(X) = X_1^{\alpha_1} \cdots X_n^{\alpha_n} u(X)$ with $u(0) \neq 0$. Then there are unitary matrices $V \in \mathrm{U}_m(\mathbb{C}[[X]])$, $U \in \mathrm{U}_d(\mathbb{C}[[X]])$ such that*

$$D(X) = V(X)^{-1}A(X)U(X)$$

is (rectangular) diagonal.

If $A = A(X) \in \mathrm{Mat}_{m,d}(\mathbb{R}[[X]])$, then U and V can be chosen real (that is orthogonal) so that $V(X)^{-1}A(X)U(X)$ is block (rectangular) diagonal with blocks as in Theorem 2.15.

Suppose in addition that the last nonzero coefficient of $\Delta_{A^*A}(X)$ is of the form $X_1^{\beta_1} \cdots X_n^{\beta_n} h(X)$ with $h(0) \neq 0$. Then, both in the real and the complex case, we may require that $V(X)^{-1}A(X)U(X)$ is (rectangular) diagonal with the entries on the diagonal in $\mathbb{R}[[X]]$, and, moreover, that those entries that are nonzero are of the form $X^\alpha a(X)$, with $a(0) > 0$, and that the exponents $\alpha \in \mathbb{N}^n$ are well-ordered.

In general, even if $d = m = 1$, one needs the extra assumption on the last nonzero coefficient of Δ_{A^*A}, see [88, Example 3.2], though this assumption is automatically satisfied if $n = 1$. It is in general not possible to have the entries of $D(X)$ nonnegative.

2.4.1. *Remarks on the proofs.* In [64] the authors first use Theorem 2.9 in order to get, after a modification of the parameter space, the eigenvalues real analytic. Then they solve linear equations describing the eigenspaces corresponding to the irreducible factors of the characteristic polynomial. This requires to monomialize the ideal defined by all the minors of the associated matrices which usually necessitates further blowings-up. A similar approach is adapted in [94]. First the eigenvalues are made analytic by blowings-up and then further blowings-up are necessary, for instance to make the coefficients of matrices and their differences normal crossing.

In [40] Grandjean shows results similar to these of [64] and [94] but by a different approach. He does not treat the eigenvalues first, but considers directly the eigenspaces defined on the complement of the discriminant locus. We refer the reader to [40] for details.

The proofs in [88] also consider the eigenvalues and eigenspaces at the same time and only the assumption that the discriminant of the reduced characteristic polynomial is normal crossings is necessary. These proofs adapt a strategy similar to the algorithm of the proof of the Abhyankar–Jung theorem of [87] and of the proof of Rellich's theorem 2.2 that we presented in Section 2.2. Given a normal matrix $A(X) \in \mathrm{Mat}_d(\mathbb{C}[[X]])$, first one subtracts from $A(X)$ the matrix $\operatorname{Tr} A$, so one can assume $\operatorname{Tr} A = 0$. This corresponds to the Tschirnhausen transformation. Then the following proposition replaces Theorem 2.12.

Proposition 2.18 ([88, Proposition 2.7])**.** *Suppose that the assumptions of Theorem 2.14 are satisfied and that, moreover, $A = (a_{ij}(X))$ is nonzero and $\operatorname{Tr} A = 0$. Then the ideal $(a_{ij})_{i,j=1,\dots,d}$ of $\mathbb{C}[[X]]$ is generated by one of a_{ij} and this generator equals a monomial in $X_1, \dots, X_n$ times a unit.*

Note that all the coefficients a_{ij} of A have the same weight, in contrast to Theorem 2.12, where suitable powers of the coefficients of P must be considered.

Finally, the following version of Hensel's lemma for normal matrices replaces the splitting lemma, Lemma 2.5. This result is a strengthened version of Cohn's version of Hensel's Lemma, see [24, Lemma 1].

Lemma 2.19 ([88, Lemma 2.1])**.** *Let $A(X) \in \mathrm{Mat}_d(\mathbb{C}[[X]])$ be a normal matrix. Assume that*

$$A(0) = \begin{pmatrix} B_1^o & 0 \\ 0 & B_2^o \end{pmatrix},$$

with $B_i^o \in \mathrm{Mat}_{d_i}(\mathbb{C})$, $d = d_1 + d_2$, such that the characteristic polynomials of B_1^o and B_2^o are coprime. Then there is a unitary matrix $U \in \mathrm{U}_d(\mathbb{C}[[X]])$, $U(0) = \mathbb{I}$, such that

$$U^{-1}AU = \begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix}, \tag{7}$$

and $B_i(0) = B_i^o$, $i = 1, 2$.

Remark 2.20. Theorems 2.14 (for $\mathbb{K} = \mathbb{C}$), 2.15 (for $\mathbb{K} = \mathbb{R}$), and 2.17 (for $\mathbb{K} = \mathbb{R}, \mathbb{C}$ remain valid if we replace $\mathbb{K}[[X]]$ by a ring $\mathbb{K}\{\!\{X\}\!\}$ satisfying the following properties:

(i) $\mathbb{K}\{\!\{X_1, \dots, X_n\}\!\}$ contains $\mathbb{K}[X_1, \dots, X_n]$,
(ii) $\mathbb{K}\{\!\{X_1, \dots, X_n\}\!\}$ is a Henselian local ring with maximal ideal $(X_1, \dots, X_n)$,
(iii) $\mathbb{K}\{\!\{X_1, \dots, X_n\}\!\} \cap (X_i)\mathbb{K}[[X_1, \dots, X_n]] = (X_i)\mathbb{K}\{\!\{X\}\!\}$ for $i = 1, \dots, n$.

Important examples of such subrings are: the algebraic power series $\mathbb{K}\langle X\rangle$, the convergent power series $\mathbb{K}\{X\}$, and the ring of germs of quasianalytic $\mathbb{K}$-valued functions over $\mathbb{R}$ (i.e. C^∞ functions satisfying (3.1) – (3.6) of [15]).

Corollary 2.21 ([88, Theorem 5.3])**.** *Let M be a manifold defined in one of the following categories:*

(i) *real analytic;*
(ii) *real Nash;*
(iii) *quasianalytic (under the asssumptions (3.1) – (3.6) of [15]).*

Let A be a matrix whose coefficients are regular functions on M (depending on the category) and let K be a compact subset of M. Then there exist a neighborhood Ω of K and the composite of a finite sequence of blowings-up with smooth centers $\pi : U \to \Omega$, such that locally on U

(a) *if A is a complex normal matrix, then $A \circ \pi$ satisfies the conclusion of Theorem 2.14;*
(b) *if A is a real normal matrix, then $A \circ \pi$ satisfies the conclusion of Theorem 2.15;*
(c) *if A is a (not necessarily square) matrix, then $A \circ \pi$ satisfies the conclusion of Theorem 2.17.*

Here one has to work with real local coordinates, since a complex change of coordinates does not commute with complex conjugation and may destroy the assumption that the matrix is normal, as seen in the following example.

Example 2.22 ([64, Example 6.1])**.** The eigenvalues of the symmetric matrix

$$A(x_1, x_2) = \begin{pmatrix} x_1^2 & x_1x_2 \\ x_1x_2 & x_2^2 \end{pmatrix},$$

are 0 and $x_1^2+x_2^2$ and $(1, x_2/x_1)$ and $(1, -x_1/x_2)$ are the corresponding eigenvectors. The discriminant of the characteristic polynomial $(x_1^2+x_2^2)^2$ is not normal crossings over the reals. It is so over the complex numbers, $(x_1+ix_2)^2(x_1-ix_2)^2$, but in the new system of coordinates (x_1+ix_2, x_1-ix_2) the matrix is not normal any longer. To diagonalize this family simultanously one has to blow-up the origin.

Remark 2.23. Many of the results presented for matrices extend (in appropriate form) to parameterized families of unbounded normal or selfadjoint operators in Hilbert space with common domain of definition and compact resolvent. See e.g. [2], [94], and references therein.

3. Differentiable roots of hyperbolic polynomials

3.1. **Bronshtein's theorem.** Let $I \subseteq \mathbb{R}$ be an open interval and

$$P_a(t)(Z) = Z^d + \sum_{j=1}^{d} a_j(t) Z^{d-j}, \quad t \in I,$$

a monic hyperbolic polynomial. It is not hard to see (e.g., [2, Theorem 4.3] and Remark 3.12 below) that if $a_j \in C^d(I)$, $1 \le j \le d$, then there exist differentiable functions $\lambda_i : I \to \mathbb{R}$, $1 \le i \le d$, which parameterize the roots of P_a, i.e.,

$$P_a(t)(Z) = \prod_{i=1}^{d} (Z - \lambda_i(t)), \quad t \in I.$$

Actually, it suffices to assume that the coefficients are of class C^p if p is the maximal multiplicity of the roots.

Bronshtein [21] proved that the roots λ_i have locally bounded derivative:

Theorem 3.1 (Bronshtein's theorem I, [21])**.** *Let $I \subseteq \mathbb{R}$ be an open interval and Y a compact Hausdorff topological space. Let $P_a(t,y)$, $(t,y) \in I \times Y$, be a monic hyperbolic polynomial of degree d such that the multiplicity of its roots does not exceed p and $\partial_t^k a_j(t,y)$, $1 \le j \le d$, $0 \le k \le p$, are continuous functions on $I \times Y$. Then, for each compact $K \subseteq I \times Y$, there is a constant $C = C(K) > 0$ such that the roots of P can be represented by functions $\lambda_i(t,y)$ that are differentiable in t and satisfy*

$$|\partial_t \lambda_i(t,y)| \le C, \quad (t,y) \in K,\ 1 \le i \le d.$$

Bronshtein's original proof is not easy to penetrate (cf. Remark 3.12). A much simpler proof was given by Parusiński and Rainer [81]. It also gives uniform bounds for the Lipschitz constant of the roots in terms of the coefficients.

Theorem 3.2 (Bronshtein's theorem II, [81])**.** *Let $I \subseteq \mathbb{R}$ be an open interval. Let $P_a(t)$, $t \in I$, be a monic hyperbolic polynomial of degree d with coefficients $a_j \in C^{d-1,1}(I)$, $1 \le j \le d$. Then any continuous root $\lambda \in C^0(I)$ of P_a is locally Lipschitz and, for any pair of relatively compact open intervals $I_0 \Subset I_1 \Subset I$,*

$$\mathrm{Lip}_{I_0}(\lambda) \le C \max_{1 \le j \le d} \|a_j\|^{1/j}_{C^{d-1,1}(\overline{I}_1)}, \tag{8}$$

with $C = C(d) \max\{\delta^{-1}, 1\}$, where $\delta := \mathrm{dist}(I_0, \mathbb{R} \setminus I_1)$.

The last claim of the above theorem, the form of the constant C, follows from the proof of [81, Theorem 2.1], see also Lemma 3.23 below. Note that the Lipschitz constant may blow up if we approach the boundary of I as for $P_a(t)(Z) = Z^d - t$

and $I = (0,1)$. We remark also that a system of continuous roots always exists, for hyperbolic polynomials one may order them for instance: $\lambda_1(t) \le \lambda_2(t) \le \ldots \le \lambda_d(t)$.

Remark 3.3. If p is the maximal multiplicity of the roots of P_a, then it is enough that $a_j \in C^{p-1,1}(I)$, $1 \le j \le d$, for any continuous root λ to be locally Lipschitz on I. The bound for the Lipschitz constant (8) of λ has to be modified; it will also depend on a quantity that measures the "uniformity" of having multiplicity at most p, namely,

$$B := \sup_{t\in I} \frac{\lambda_d(t) - \lambda_1(t)}{\min_{1\le i\le d-p}(\lambda_{i+p}(t) - \lambda_i(t))},$$

where $\lambda_1(t) \le \lambda_2(t) \le \ldots \le \lambda_d(t)$, $t \in I$, is an increasing enumeration of the roots. See [81, Section 4.7] for details.

We will give a full proof of Theorem 3.2 in Section 3.2. A multiparameter version follows immediately (because a function defined on a box that is Lipschitz with respect to each variable is Lipschitz).

Theorem 3.4 ([81, Theorem 2.2])**.** *Let $U \subseteq \mathbb{R}^n$ be open. Let $P_a(x)$, $x \in U$, be a $C^{d-1,1}$ family of monic hyperbolic polynomials of degree d. Then any continuous root $\lambda \in C^0(U)$ of P_a is locally Lipschitz and, for any pair of relatively compact open subsets $U_0 \Subset U_1 \Subset U$,*

$$\operatorname{Lip}_{U_0}(\lambda) \le C(d,n,U_0,U_1) \max_{1\le i\le d} \|a_i\|^{1/i}_{C^{d-1,1}(\overline{U}_1)}. \tag{9}$$

Recall that $\operatorname{Hyp}(d)$ denotes the space of monic hyperbolic polynomials of degree d. Then $\operatorname{Hyp}(d)$ can be identified with a closed semialgebraic subset of $\mathbb{R}^d$; see Section 8. Consider the map $\lambda^\uparrow : \operatorname{Hyp}(d) \to \mathbb{R}^d$ which assigns to P_a its increasingly ordered roots. Then $\lambda^\uparrow$ is continuous.

Corollary 3.5. *Let $U \subseteq \mathbb{R}^n$ be open. Then the push forward*

$$(\lambda^\uparrow)_* : C^{d-1,1}(U, \operatorname{Hyp}(d)) \to C^{0,1}(U, \mathbb{R}^d), \quad P_a \mapsto \lambda^\uparrow \circ P_a,$$

is bounded.

It is natural to ask if the map $(\lambda^\uparrow)_*$ is continuous. The answer is no:

Example 3.6. Let $f(t) := t^2$ and $f_n(t) := t^2 + 1/n^2$, $n \ge 1$. Then, for all $k \in \mathbb{N}$ and each bounded open interval $I \subseteq \mathbb{R}$, $\|f - f_n\|_{C^k(\overline{I})} = 1/n^2 \to 0$ as $n \to \infty$. Let λ and λ_n be the positive square roots of f and f_n, respectively: $\lambda(t) := |t|$ and $\lambda_n(t) := \sqrt{t^2 + 1/n^2}$. Then, for each bounded open interval $I \subseteq \mathbb{R}$ that contains 0,

$$\begin{aligned}\operatorname{Lip}_I(\lambda - \lambda_n) &\ge \sup_{0<t\in I} \Big| \frac{(\lambda(t) - \lambda_n(t)) - (\lambda(0) - \lambda_n(0))}{t} \Big| \\ &= \sup_{0<t\in I} \Big| \frac{t - \sqrt{t^2 + \frac{1}{n^2}} + \frac{1}{n}}{t} \Big| \ge \Big| \frac{\frac{1}{n} - \sqrt{\frac{1}{n^2} + \frac{1}{n^2}} + \frac{1}{n}}{\frac{1}{n}} \Big| = 2 - \sqrt{2},\end{aligned}$$

for large enough n. Consequently, the map $(\lambda^\uparrow)_* : C^k(\mathbb{R}, \operatorname{Hyp}(2)) \to C^{0,1}(\mathbb{R}, \mathbb{R}^2)$ (where $k \ge 2$) is not continuous.

But $(\lambda^\uparrow)_* : C^{1,1}(\mathbb{R}, \operatorname{Hyp}(2)) \to W^{1,p}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^2)$ is continuous for each $1 \le p < \infty$:

Remark 3.7. Let $f \in C^{1,1}(\mathbb{R},[0,\infty))$ and $f_n \in C^{1,1}(\mathbb{R},[0,\infty))$, $n \geq 1$, such that $\|f - f_n\|_{C^{1,1}(\bar{I})} \to 0$ as $n \to \infty$, for each bounded open interval $I \subseteq \mathbb{R}$. Let λ and λ_n be the nonnegative square roots of f and f_n, respectively. They are locally Lipschitz (by Theorem 3.2) and thus differentiable almost everywhere. Then

$$(\lambda - \lambda_n)(\lambda + \lambda_n) = f - f_n,$$
$$(\lambda' - \lambda_n')(\lambda + \lambda_n) + (\lambda - \lambda_n)(\lambda' + \lambda_n') = f' - f_n',$$

and, where $\lambda + \lambda_n \neq 0$,

$$\lambda' - \lambda_n' = \frac{f' - f_n'}{\lambda + \lambda_n} - \frac{\lambda - \lambda_n}{\lambda + \lambda_n}(\lambda' + \lambda_n').$$

Note that at points t, where $\lambda(t) + \lambda_n(t) = 0$, we have $\lambda(t) = \lambda_n(t) = 0$ as well as $f(t) = f_n(t) = f'(t) = f_n'(t) = 0$. Let $I \subseteq \mathbb{R}$ be a bounded open interval and $1 \leq p < \infty$. By Glaeser's inequality (16), we have on I,

$$\Big|\frac{f' - f_n'}{\lambda + \lambda_n}\Big|^p \leq \Big(\frac{|f'| + |f_n'|}{\lambda + \lambda_n}\Big)^p \leq \Big(\frac{\sqrt{2C_1}(\lambda + \lambda_n)}{\lambda + \lambda_n}\Big)^p = (2C_1)^{p/2},$$

where $C_1 := \max\{\|((\psi f)''\|_{L^\infty(\mathbb{R})}, \sup_n \|(\psi f_n)''\|_{L^\infty(\mathbb{R})}\}$ and $\psi \geq 0$ is a C^∞ function which equals 1 in a neighborhood of I and has compact support (say in a 1-neighborhood of I). Thus,

$$\Big\|\frac{f' - f_n'}{\lambda + \lambda_n}\Big\|_{L^p(I)}^p \to 0 \quad \text{as } n \to \infty,$$

by the dominated convergence theorem. Moreover, on I,

$$\Big|\frac{\lambda - \lambda_n}{\lambda + \lambda_n}(\lambda' + \lambda_n')\Big|^p \leq (2C_2)^p \Big|\frac{\lambda - \lambda_n}{\lambda + \lambda_n}\Big|^p \leq (2C_2)^p,$$

where $C_2 := \max\{\operatorname{Lip}_I(\lambda), \sup_n \operatorname{Lip}_I(\lambda_n)\}$ (which is finite by Theorem 3.2). So

$$\Big\|\frac{\lambda - \lambda_n}{\lambda + \lambda_n}(\lambda' + \lambda_n')\Big\|_{L^p(I)}^p \to 0 \quad \text{as } n \to \infty,$$

again by the dominated convergence theorem. Thus,

$$\|\lambda' - \lambda_n'\|_{L^p(I)} \leq \Big\|\frac{f' - f_n'}{\lambda + \lambda_n}\Big\|_{L^p(I)} + \Big\|\frac{\lambda - \lambda_n}{\lambda + \lambda_n}(\lambda' + \lambda_n')\Big\|_{L^p(I)} \to 0,$$

as $n \to \infty$. We conclude that the map $(\lambda^\uparrow)_* : C^{1,1}(\mathbb{R}, \operatorname{Hyp}(2)) \to W^{1,p}_{\text{loc}}(\mathbb{R}, \mathbb{R}^2)$ is continuous for each $1 \leq p < \infty$.

Open Problem 3.8. *Is this also true for $d \geq 3$?*

It turns out that the derivatives of a differentiable choice of the roots not only are bounded but they are continuous, provided that the coefficients are C^p, where p is the maximal multiplicity of the roots. This was first observed by Colombini, Orrú, and Pernazza [27] and a short proof was given by Parusiński and Rainer [81].

Theorem 3.9 ([81, Theorem 2.4])**.** *Let $P_a(t)$, $t \in I$, be a C^p curve of monic hyperbolic polynomials, where p is the maximal multiplicity of the roots. Then:*

(1) *Any continuous root λ of P_a has left- and right-sided derivatives $\lambda'^{\pm}$ at each $t \in I$.*
(2) *λ'^- and λ'^+ are continuous in the following sense: for every $t_0 \in I$, $\lim_{t \to t_0^-} \lambda'^{\pm}(t) = \lambda'^-(t_0)$ and $\lim_{t \to t_0^+} \lambda'^{\pm}(t) = \lambda'^+(t_0)$.*
(3) *There exist differentiable systems of the roots, any differentiable root is C^1.*

If the coefficients in Theorem 3.9 are of class C^{2p}, then there are even C^1 systems of the roots such that their derivatives of second order exist everywhere. This is due to Colombini, Orrú, and Pernazza [27] following the strategy of Kriegl, Losik, and Michor [55] who needed C^{3p} coefficients. In general, these results cannot be improved, even for C^∞ coefficients.

Example 3.10. (1) The function $f : \mathbb{R} \to [0, \infty)$ given by

$$f(t) := \begin{cases} e^{-1/|t|}(\sin^2(\pi/|t|) + e^{-1/t^2}) & \text{if } t \neq 0, \\ 0 & \text{if } t \neq 0, \end{cases}$$

is C^∞, but no solution of $Z^2 = f(t)$ is $C^{1,\alpha}$ for any $\alpha > 0$ (cf. [18]). It was shown in [18] that, given any modulus of continuity ω, there exists a C^∞ function on $\mathbb{R}$ (depending on ω) such that no solution of $Z^2 = f(t)$ is $C^{1,\omega}$.

(2) A nonnegative $C^{1,1}$ function g such that no solution of $Z^2 = g(t)$ is differentiable at $t = 0$ is

$$g(t) := \begin{cases} t^2 \sin^2(\log |t|) & \text{if } t \neq 0, \\ 0 & \text{if } t = 0. \end{cases}$$

(3) Replacing t^2 by t^4 in the definition of g we get a $C^{3,1}$ function h such that no solution of $Z^2 = h(t)$ is twice differentiable at $t = 0$.

(4) Clearly, the solutions of our equation cannot in general be locally Lipschitz if the right-hand side is not $C^{1,1}$, e.g., $Z^2 = t^{1+\alpha}$, where $\alpha \in (0, 1)$.

We refer to [27] for other illuminating examples (including hyperbolic polynomials of degree ≥ 3).

Wakabayashi [114] gave a different (complex analytic) proof of a more general Hölder version of Bronshtein's theorem which had been announced by Ohya and Tarama [80]. A proof based on Bronshtein's original method can be found in Tarama [109].

Theorem 3.11 (Bronshtein's theorem III, [109, Theorem 1.4])**.** *Let $I \subseteq \mathbb{R}$ be an open bounded interval. Let $P_a(t)$, $t \in I$, be a monic hyperbolic polynomial of degree d such that the multiplicity of its roots does not exceed $p \geq 2$ and $a_j \in C^{k,\alpha}(I)$, $1 \leq j \leq d$. Then each continuous root λ of $P_a(t)$ is locally Hölder continuous of index*

$$\beta := \min\left\{1, \tfrac{k+\alpha}{p}\right\}$$

on I. Moreover, if the coefficients form a bounded subset of $C^{k,\alpha}(\overline{I})$, then the continuous roots form a bounded subset of $C^{0,\beta}(\overline{J})$, for each relatively compact open subinterval $J \Subset I$, as long as the polynomial is hyperbolic with the multiplicity of the roots uniformly not exceeding p, i.e., B defined in Remark 3.3 is finite.

The main motivation of Bronshtein for proving his theorem was to show Gevrey well-posedness of the Cauchy problem with multiple characteristics; see [22]. To this end, he deduces from the Lipschitz continuity of the roots of $P_a(t)$ the bound

$$\Big|\frac{\partial_t P_a(t)(z)}{P_a(t)(z)}\Big| \lesssim |\mathrm{Im}(z)|^{-1}, \quad 0 < |\mathrm{Im}(z)| \leq 1.$$

In the setting of Theorem 3.11, Tarama shows in [109] that

$$\Big|\frac{\partial_t^j P_a(t)(z)}{P_a(t)(z)}\Big| \lesssim |\mathrm{Im}(z)|^{-j \max\{1, 1/\beta\}}, \tag{10}$$

for $1 \le j \le \min\{p-1, k\}$, $0 < |\mathrm{Im}(z)| \le 1$, and t in a compact subinterval of I. It is also proved in [109] that the statements of Theorem 3.11 and (10) are equivalent.

Remark 3.12. Bronshtein's original proof of Theorem 3.1 rests on a polynomial equation for the derivatives of the roots of P_a, see (13) below. Suppose that λ_0 is a q-fold root of $P_a(t_0)$ and consider the polynomial

$$P_a(t)(Z+\lambda_0) = A_0(t)Z^q + A_1(t)Z^{q-1} + \cdots + A_q(t) + R(t,Z), \tag{11}$$

where $R(t,Z)$ is a polynomial in Z divisible by Z^{q+1} (possibly identically zero). Note that $A_j(t) = \frac{1}{(q-j)!}\partial_Z^{q-j} P_a(t)(\lambda_0)$ and set $a_j^i := \partial_t^i A_j(t_0)$. Then $a_0^0 \neq 0$, while $a_j^0 = 0$ for $1 \le j \le q$, since λ_0 is a q-fold root of $P_a(t_0)$. Using that P_a is hyperbolic and that hyperbolicity is preserved by differentiation with respect to Z, Bronshtein shows that $a_j^i = 0$ for $0 \le i \le j-1$ and $1 \le j \le q$, or equivalently, that $A_j(t) = (t-t_0)^j \hat{A}_j(t)$, where $\hat{A}_j$ is a continuous function.

Consequently, the change of variables $Z = (t-t_0)W$ in (11) leads to

$$\begin{aligned}(t-t_0)^{-q}P_a(t)((t-t_0)W+\lambda_0) = a_0^0\, W^q + \tfrac{1}{1!}a_1^1\, W^{q-1} + \cdots + \tfrac{1}{q!}a_q^q \\ + \sum_{k=0}^{d} B_k(t)W^{d-k},\end{aligned} \tag{12}$$

where B_k are continuous functions with $B_k(t_0) = 0$. It follows that P_a has q roots of the form $\lambda_j(t) = \lambda_0 + (t-t_0)\mu_j(t)$, where the μ_j are the roots of (12) which are continuous at t_0. That means that all the roots of P_a having the value λ_0 at t_0 are differentiable at t_0 and the derivatives are the solutions of

$$a_0^0\, W^q + \tfrac{1}{1!}a_1^1\, W^{q-1} + \cdots + \tfrac{1}{q!}a_q^q = 0. \tag{13}$$

Now, for contradiction, Bronshtein assumes that there is a sequence $(t_n, y_n) \to (t_0, y_0)$ such that $\lambda(t_n, y_n) \to \lambda(t_0, y_0)$ and $|\partial_t \lambda(t_n, y_n)| \to \infty$, and that the multiplicity of $\lambda(t_n, y_n)$ as a root of $P_a(t_n, y_n)$ is q for all $n \ge 1$. Then $\partial_t \lambda(t_n, y_n)$ satisfies an equation of type (13) which is found as above, where $\lambda(t_n, y_n)$ plays the role of λ_0; in particular, the coefficients depend on n. It remains to show that the coefficients of this equation, after division by the coefficient of W^q, are bounded in n. This is the hard part of Bronshtein's proof (see [21, Lemma 4 and 4']).

3.2. **Towards a proof of Bronshtein's theorem.** We will give a simple proof of Theorem 3.2 which is based on [81]. It requires some preparation.

3.2.1. *Splitting.* We proceed as in Section 2.2.2, but replace the analytic function θ by $|\tilde{a}_2|^{1/2}$. Let $P_{\tilde{a}} \in \mathrm{Hyp}_T(d)$ be such that $\tilde{a} \neq 0$. Then the polynomial

$$Q_{\underline{a}}(Z) := |\tilde{a}_2|^{-d/2} P_{\tilde{a}}(|\tilde{a}_2|^{1/2} Z) = Z^d - Z^{d-2} + \sum_{j=3}^{d} |\tilde{a}_2|^{-j/2}\tilde{a}_j Z^{d-j}$$

belongs to $\mathrm{Hyp}_T^0(d) = \{P_{\tilde{a}} \in \mathrm{Hyp}_T(d) : \tilde{a}_2 = -1\}$. By Lemma 2.5, we have

$$Q_{\underline{a}} = Q_{\underline{b}} Q_{\underline{c}},$$

on some open ball $B(P_{\tilde{a}}, r) \subseteq \mathbb{R}^d$ such that $\deg Q_{\underline{b}} < d$, $\deg Q_{\underline{c}} < d$, and

$$\underline{b}_i = \psi_i(|\tilde{a}_2|^{-3/2}\tilde{a}_3, \ldots, |\tilde{a}_2|^{-d/2}\tilde{a}_d), \quad i = 1, \ldots, \deg Q_{\underline{b}},$$

where ψ_i are real analytic functions; likewise for $\underline{c}_i$. If $Q_{\underline{a}}$ is hyperbolic, then also $Q_{\underline{b}}$ and $Q_{\underline{c}}$ are hyperbolic; we restrict our attention to the set $B(P_{\tilde{a}}, r) \cap \operatorname{Hyp}_T(d)$. This induces a splitting

$$P_{\tilde{a}} = P_b P_c, \qquad \text{on } B(P_{\tilde{a}}, r),$$

where

$$b_i = |\tilde{a}_2|^{i/2} \psi_i(|\tilde{a}_2|^{-3/2}\tilde{a}_3, \dots, |\tilde{a}_2|^{-d/2}\tilde{a}_d), \quad i = 1, \dots, \deg P_b. \tag{14}$$

The coefficients $\tilde{b}_i$ of $P_{\tilde{b}}$ resulting from P_b by the Tschirnhausen transformation have an analogous representation (in view of (1))

$$\tilde{b}_i = |\tilde{a}_2|^{i/2} \tilde{\psi}_i(|\tilde{a}_2|^{-3/2}\tilde{a}_3, \dots, |\tilde{a}_2|^{-d/2}\tilde{a}_d), \quad i = 1, \dots, \deg P_b. \tag{15}$$

Shrinking $r > 0$ slightly, we may assume that all partial derivatives of ψ_i and $\tilde{\psi}_i$ of all orders are bounded on $B(P_{\tilde{a}}, r)$.

Lemma 3.13. *In this situation we have* $|\tilde{b}_2| \le 4\,|\tilde{a}_2|$.

Proof. Let $\lambda_1, \dots, \lambda_d$ be the roots of $P_{\tilde{a}}$ such that the first k of them, i.e., $\lambda_1, \dots, \lambda_k$, are the roots of P_b. Then $|b_1| \le \sum_{j=1}^k |\lambda_j| \le \sqrt{k}\,\big(\sum_{j=1}^k \lambda_j^2\big)^{1/2}$ and thus

$$\begin{aligned} 2\,|\tilde{b}_2| &= \sum_{j=1}^k \Big(\lambda_j + \frac{b_1}{k}\Big)^2 = \sum_{j=1}^k \Big(\lambda_j^2 + \frac{b_1^2}{k^2} + \frac{2}{k}\lambda_j b_1\Big) \\ &= \sum_{j=1}^k \lambda_j^2 + \frac{b_1^2}{k^2}\cdot k + \frac{2}{k} b_1 \sum_{j=1}^k \lambda_j \le (1+1+2)\sum_{j=1}^k \lambda_j^2 \le 8\,|\tilde{a}_2|, \end{aligned}$$

since $\sum_{j=1}^d \lambda_j^2 = -2\tilde{a}_2$. □

3.2.2. *Glaeser's inequality.* A classical inequality used by Glaeser in [39] (and attributed by him to Malgrange) is the following: for nonnegative C^1 functions f on $\mathbb{R}$ with $f'' \in L^\infty(\mathbb{R})$ we have

$$f'(t)^2 \le 2 f(t) \|f''\|_{L^\infty(\mathbb{R})}, \quad t \in \mathbb{R}. \tag{16}$$

Note that this shows Bronshtein's theorem in the simplest nontrivial case. We shall need a local version. For $t_0 \in \mathbb{R}$ and $r > 0$, let $I(t_0, r)$ denote the open interval centered at t_0 with radius r,

$$I(t_0, r) := \{t \in \mathbb{R} : |t - t_0| < r\}.$$

Lemma 3.14. *Let $I \subseteq \mathbb{R}$ be an open bounded interval. Let $f \in C^{1,1}(\overline{I})$ satisfy $f \ge 0$ or $f \le 0$ on I. Let $M > 0$ and assume that $t_0 \in I$, $f(t_0) \ne 0$, and $I_0 := I(t_0, M^{-1}|f(t_0)|^{1/2}) \subseteq I$. Then*

$$|f'(t_0)| \le (M + M^{-1} \operatorname{Lip}_{I_0}(f'))|f(t_0)|^{1/2}.$$

If additionally $\operatorname{Lip}_{I_0}(f') \le M^2$, *then*

$$|f'(t_0)| \le 2M\,|f(t_0)|^{1/2}.$$

Note that if $f(t_0) = 0$ also $f'(t_0) = 0$.

Proof. Suppose that $f \ge 0$; otherwise consider $-f$. Thus $f(t_0) > 0$ and

$$0 \le f(t_0 + h) = f(t_0) + f'(t_0)h + \int_0^1 f'(t_0 + hs) - f'(t_0)\,ds \cdot h.$$

The assertion follows from setting $h := \pm M^{-1}|f(t_0)|^{1/2}$. □

3.2.3. *Interpolation estimates.*

Lemma 3.15. *Let $T(x) = a_0 + a_1 x + \cdots + a_m x^m \in \mathbb{C}[x]$ satisfy $|T(x)| \le A$ for $x \in [0, B] \subseteq \mathbb{R}$. Then*

$$|a_j| \le C(m)\, AB^{-j}, \quad 0 \le j \le m.$$

Proof. It suffices to assume $A = B = 1$; for the general case consider $A^{-1}T(By)$. Noting that $(T(x_0), \dots, T(x_m))^t = V(a_0, \dots, a_m)^t$ and hence $(a_0, \dots, a_m)^t = V^{-1}(T(x_0), \dots T(x_m))^t$, where V is the Vandermonde matrix of the equidistant points $0 = x_0 < x_1 < \cdots < x_m = 1$, the statement follows easily. □

Lemma 3.16. *Let $f \in C^{m,1}(\overline{I})$, where $I \subseteq \mathbb{R}$ is a bounded open interval. Then*

$$|f^{(k)}(t)| \le C(m)\, |I|^{-k} \big(\|f\|_{L^\infty(I)} + \mathrm{Lip}_I(f^{(m)}) |I|^{m+1} \big), \quad t \in I,\ 1 \le k \le m.$$

Proof. Let $t \in I$. Then $[t, t + |I|/2)$ or $(t - |I|/2, t]$ is contained in I. Let s be a point in the respective interval. By Taylor's formula,

$$\begin{aligned} &\Big| \sum_{k=0}^{m} \frac{f^{(k)}(t)}{k!} (s-t)^k \Big| \\ &\quad = \Big| f(s) - (s-t)^m \int_0^1 \frac{(1-u)^{m-1}}{(m-1)!} (f^{(m)}(t + u(s-t)) - f^{(m)}(t))\, du \Big| \\ &\quad \le \|f\|_{L^\infty(I)} + |I|^{m+1} \mathrm{Lip}_I(f^{(m)}). \end{aligned}$$

Now apply Lemma 3.15. □

3.2.4. *The key argument.*

Definition 3.17. Let $I_1 \subseteq \mathbb{R}$ be an open bounded interval and $I_0 \Subset I_1$ a relatively compact open subinterval. Let $P_{\tilde{a}}(t)$, $t \in I_1$, be a monic hyperbolic polynomial of degree d in Tschirnhausen form with coefficients $\tilde{a}_j \in C^{d-1,1}(\overline{I}_1)$, $j = 2, \dots, d$. Let $A > 0$ be a constant. We say that $(P_{\tilde{a}}, I_1, I_0, A)$ is *admissible* if for every $t_0 \in I_0 \setminus \{t : \tilde{a}_2(t) = 0\}$

$$I_A(t_0) := I(t_0, A^{-1} |\tilde{a}_2(t_0)|^{1/2}) \subseteq I_1, \tag{17}$$

$$\frac{1}{2} \le \frac{\tilde{a}_2(t)}{\tilde{a}_2(t_0)} \le 2, \quad t \in I_A(t_0), \tag{18}$$

$$|\tilde{a}_j^{(k)}(t)| \le A^k\, |\tilde{a}_2(t_0)|^{(j-k)/2}, \quad 2 \le j \le d,\ 1 \le k \le d-1,\ t \in I_A(t_0), \tag{19}$$

$$\mathrm{Lip}_{I_A(t_0)}(\tilde{a}_j^{(d-1)}) \le A^d\, |\tilde{a}_2(t_0)|^{(j-d)/2}, \quad 2 \le j \le d. \tag{20}$$

The following lemma is an easy exercise.

Lemma 3.18. *Let $(P_{\tilde{a}}, I_1, I_0, A)$ be admissible. Then the functions $\underline{a}_j := |\tilde{a}_2|^{-j/2} \tilde{a}_j$, $2 \le j \le d$, are well-defined on $I_A(t_0)$ and satisfy*

$$|\underline{a}_j^{(k)}(t)| \le A^k\, |\tilde{a}_2(t_0)|^{-k/2}, \quad 2 \le j \le d,\ 1 \le k \le d-1,\ t \in I_A(t_0), \tag{21}$$

$$\mathrm{Lip}_{I_A(t_0)}(\underline{a}_j^{(d-1)}) \le A^d\, |\tilde{a}_2(t_0)|^{-d/2}, \quad 2 \le j \le d. \tag{22}$$

There is some redundancy in the conditions (17)–(20), up to multiplying A by some constant $C(d) \ge 1$:

Lemma 3.19. *Suppose that (17), (19) for $k \ge j$, $2 \le j \le d$, and (20) hold. Then there is a constant $C(d) \ge 1$ such that $(P_{\tilde{a}}, I_1, I_0, C(d)A)$ is admissible.*

Proof. By (19) for $j = k = 2$, $\mathrm{Lip}_{I_A(t_0)}(\tilde{a}_2') \le A^2$. Lemma 3.14 and (17) imply

$$|\tilde{a}_2'(t_0)| \le 2A\,|\tilde{a}_2(t_0)|^{1/2}.$$

Thus, for $t \in I_{6A}(t_0)$,

$$\begin{aligned}|\tilde{a}_2(t) - \tilde{a}_2(t_0)| &\le |\tilde{a}_2'(t_0)||t-t_0| + \mathrm{Lip}_{I_A(t_0)}(\tilde{a}_2')\,|t-t_0|^2\\ &\le \frac{1}{3}\,|\tilde{a}_2(t_0)| + \frac{1}{36}\,|\tilde{a}_2(t_0)| < \frac{1}{2}\,|\tilde{a}_2(t_0)|,\end{aligned}$$

which gives (18). That (19) also holds for $k < j$ follows from Lemma 3.16 applied to $f = \tilde{a}_j$ and $m = j-1$, together with (18) and Lemma 2.4. □

Proposition 3.20. *Let $(P_{\tilde{a}}, I_1, I_0, A)$ be admissible and $t_0 \in I_0 \setminus \{t : \tilde{a}_2(t) = 0\}$. Then there exist a constant $C(d) > 1$ and open bounded intervals $J_1 \Supset J_0 \ni t_0$, J_0 relatively compact in J_1, such that the following holds. We have*

$$P_{\tilde{a}} = P_b P_c, \qquad \text{on } J_1,$$

where P_b and P_c are monic hyperbolic polynomials of degree $< d$ and with coefficients in $C^{d-1,1}(\overline{J}_1)$, and, after Tschirnhausen transformation, $(P_{\tilde{b}}, J_1, J_0, C(d)A)$ and $(P_{\tilde{c}}, J_1, J_0, C(d)A)$ are admissible.

Proof. Consider

$$\underline{a} := (-1, \underline{a}_3, \dots, \underline{a}_d) : I_A(t_0) \to \mathbb{R}^{d-1},$$

which defines a continuous bounded curve, by Lemma 2.4. By Lemma 3.18,

$$|\underline{a}'(t)| \le A\,|\tilde{a}_2(t_0)|^{-1/2}, \quad t \in I_A(t_0). \tag{23}$$

Choose a finite cover of $\mathrm{Hyp}_T^0(d)$ by balls $B_1, \dots, B_s$ such that on each B_i we have a splitting of $P_{\tilde{a}}$ (cf. Section 3.2.1). There exists $r_1 \in (0,1)$ such that for any $p \in \mathrm{Hyp}_T^0(d)$ there is $i \in \{1, \dots, s\}$ such that $B(p, r_1) \subseteq B_i$ (note that $2r_1$ is a Lebesgue number of the cover $B_1, \dots, B_s$). Set

$$J_1 := I_{A/r_1}(t_0).$$

Then $J_1 \subseteq \underline{a}^{-1}(B(\underline{a}(t_0), r_1))$ so that

$$P_{\tilde{a}} = P_b P_c, \qquad \text{on } J_1.$$

Fix $r_0 < r_1$ and set $J_0 := I_{A/r_0}(t_0)$. The coefficients of $P_{\tilde{b}}$ (after Tschirnhausen transformation) are given by the formulas (15).

Let us show that $(P_{\tilde{b}}, J_1, J_0, B)$ is admissible, where $B = C(d)A$ for a suitable constant $C(d)$. If B is a constant satisfying

$$B \ge \frac{2\sqrt{2}\,A}{r_1 - r_0},$$

then, for each $t_1 \in J_0 \setminus \{t : \tilde{b}_2(t) = 0\}$,

$$B^{-1}|\tilde{b}_2(t_1)|^{1/2} \le \frac{r_1 - r_0}{A}\,|\tilde{a}_2(t_0)|^{1/2},$$

by Lemma 3.13 and (18). This implies

$$J_B(t_1) := I(t_1, B^{-1}|\tilde{b}_2(t_1)|^{1/2}) \subseteq J_1,$$

by the definition of J_1 and J_0. Note that r_1 and r_0 can be chosen in a universal way so that we subsume the dependence of B on them in $C(d)$.

Next we claim that, on J_1,

$$|\partial_t^k[\tilde{\psi}_i \circ (\underline{a}_3, \dots, \underline{a}_d)]| \le C(d)A^k\,|\tilde{a}_2(t_0)|^{-k/2}, \quad 0 \le k \le d-1,$$
$$\mathrm{Lip}_{J_1}(\partial_t^{d-1}[\tilde{\psi}_i \circ (\underline{a}_3, \dots, \underline{a}_d)]) \le C(d)A^d\,|\tilde{a}_2(t_0)|^{-d/2}.$$

Recall that the real analytic functions $\tilde{\psi}_i$ and its partial derivatives of all orders are bounded by universal constants. Thus the claim is obvious for $k = 0$. We have

$$\partial_t[\tilde{\psi}_i \circ (\underline{a}_3, \dots, \underline{a}_d)] = d\tilde{\psi}_i(\underline{a})(\underline{a}')$$

so that the claim for $k = 1$ follows from (23). For $2 \le k \le d-1$, the claim follows from differentiating this equation and using (21). In a similar way, using also (22), one gets the estimate for $\mathrm{Lip}_{J_1}(\partial_t^{d-1}[\tilde{\psi}_i \circ (\underline{a}_3, \dots, \underline{a}_d)])$.

Now it is easy to conclude (from the formulas (15))

$$|\tilde{b}_i^{(k)}(t)| \le C(d)A^k\,|\tilde{a}_2(t_0)|^{(i-k)/2}, \quad t \in J_1,\ 1 \le k \le d-1,$$
$$\mathrm{Lip}_{J_1}(\tilde{b}_i^{(d-1)}) \le C(d)A^d\,|\tilde{a}_2(t_0)|^{(i-d)/2},$$

for all $2 \le i \le \deg P_b$. By Lemma 3.13 and (18), $|\tilde{a}_2(t_0)|^{-1} \le 2\,|\tilde{a}_2(t_1)|^{-1} \le 8\,|\tilde{b}_2(t_1)|^{-1}$ and thus

$$|\tilde{b}_i^{(k)}(t)| \le C(d)A^k\,|\tilde{b}_2(t_1)|^{(i-k)/2}, \quad t \in J_1,\ i \le k \le d-1,$$
$$\mathrm{Lip}_{J_1}(\tilde{b}_i^{(d-1)}) \le C(d)A^d\,|\tilde{b}_2(t_1)|^{(i-d)/2},$$

for all $2 \le i \le \deg P_b$. Thus we may conclude that $(P_{\tilde{b}}, J_1, J_0, B)$ is admissible from Lemma 3.19. □

Remark 3.21. We have the same estimates for b_i instead of $\tilde{b}_i$.

Theorem 3.22. *Let $(P_{\tilde{a}}, I_1, I_0, A)$ be admissible. Let $\lambda : I_1 \to \mathbb{R}$ be a continuous root of $P_{\tilde{a}}$. Then λ is Lipschitz on I_0 with $\mathrm{Lip}_{I_0}(\lambda) \le C(d)A$.*

Proof. We proceed by induction on the degree. The only monic polynomial of degree 1 in Tschirnhausen form is Z; so there is nothing to prove. So assume that $\deg P_{\tilde{a}} \ge 2$. Fix $t_0 \in I_0 \setminus \{t : \tilde{a}_2(t) = 0\}$. By Proposition 3.20, we may assume that on a neighborhood J_1 of t_0 we have

$$\lambda(t) = -\frac{b_1(t)}{\deg P_b} + \mu(t),$$

where μ is a continuous root of $P_{\tilde{b}}$ and $(P_{\tilde{b}}, J_1, J_0, C(d)A)$ is admissible. By the induction hypothesis, μ is Lipschitz on J_0 with $\mathrm{Lip}_{J_0}(\mu) \le C(d)A$. Thus λ is Lipschitz on J_0 with $\mathrm{Lip}_{J_0}(\lambda) \le C(d)A$, in view of Remark 3.21. It is an easy exercise to show that these local uniform Lipschitz bounds on $I_0 \setminus \{t : \tilde{a}_2(t) = 0\}$ imply that $\mathrm{Lip}_{I_0}(\lambda) \le C(d)A$, since λ vanishes on the set $\{t : \tilde{a}_2(t) = 0\}$. □

3.2.5. *Proof of Bronshtein's theorem.* Before we start with the proof, we observe in the following lemma that the assumptions of Theorem 3.2 (after Tschirnhausen transformation) yield an admissible quadruple.

Lemma 3.23. *Let $I_1 \subseteq \mathbb{R}$ be a bounded open interval and $P_{\tilde{a}}$ a monic hyperbolic polynomial of degree d in Tschirnhausen form with coefficients $\tilde{a}_j \in C^{d-1,1}(\bar{I}_1)$, $2 \le j \le d$. If $I_0 \Subset I_1$ is an open subinterval, relatively compact in I_1, then $(P_{\tilde{a}}, I_1, I_0, A)$ is admissible with*

$$A := 6\max\{A_1, A_2\}, \tag{24}$$

where

$$A_1 := \max\left\{\delta^{-1}\|\tilde a_2\|^{1/2}_{L^\infty(I_1)}, (\operatorname{Lip}_{I_1}(\tilde a_2'))^{1/2}\right\},$$
$$A_2 := \max_{2\le j\le d}\left\{\operatorname{Lip}_{I_1}(\tilde a_j^{(d-1)})\,\|\tilde a_2\|^{(d-j)/2}_{L^\infty(I_1)}\right\}^{1/d},$$

and $\delta := \operatorname{dist}(I_0, \mathbb{R}\setminus I_1)$.

Proof. Let $t_0 \in I_0\setminus\{t : \tilde a_2(t) = 0\}$. We have $I_{A_1}(t_0) \subseteq I_1$, i.e., (17). By Lemma 3.14,

$$|\tilde a_2'(t_0)| \le 2A_1\,|\tilde a_2(t_0)|^{1/2}.$$

This implies (as in the proof of Lemma 3.19) that (18) holds for $t \in I_{6A_1}(t_0)$. Now (20) is clear from the definition of A_2. Finally, Lemma 2.4 and Lemma 3.16 imply (19), for $t \in I_A(t_0)$. □

Now let $P_{\tilde a}$ be a monic hyperbolic polynomial of degree d in Tschirnhausen form with coefficients $\tilde a_j \in C^{d-1,1}(I)$, $2 \le j \le d$. Fix relatively compact open subintervals $I_0 \Subset I_1 \Subset I$. By Lemma 3.23 and Theorem 3.22, any continuous root λ of $P_{\tilde a}$ is Lipschitz on I_0 with $\operatorname{Lip}_{I_0}(\lambda) \le C(d)A$, where A is defined in (24). This implies

$$\operatorname{Lip}_{I_0}(\lambda) \le C(d)\,\max\{\delta^{-1}, 1\}\,\max_{2\le j\le d}\|\tilde a_j\|^{1/j}_{C^{d-1,1}(\bar I_1)}.$$

If P_a is not necessarily in Tschirnhausen form, we apply the Tschirnhausen transformation $P_a \rightsquigarrow P_{\tilde a}$. Then $\tilde a_j$ is a weighted homogeneous polynomial of degree j in $a_1, \dots, a_d$, where a_j has the weight j; see (1). The roots are shifted by a_1/d. Thus any continuous root λ of P_a is Lipschitz on I_0 and

$$\operatorname{Lip}_{I_0}(\lambda) \le C(d)\,\max\{\delta^{-1}, 1\}\,\max_{1\le j\le d}\|a_j\|^{1/j}_{C^{d-1,1}(\bar I_1)}.$$

This ends the proof of Theorem 3.2.

3.3. **Sufficient conditions for C^p roots.** Let us analyze what causes the loss of regularity and give several sufficient conditions for better regularity of the roots.

3.3.1. *The effect of positive local minima.* A finer analysis of Example 3.10(1) indicates that the small positive local minima of the function f prevent the solutions of $Z^2 = f$ from being $C^{1,\alpha}$. This is confirmed by the following result of Bony, Broglia, Colombini, and Pernazza [18].

Theorem 3.24 ([18, Theorem 3.5])**.** *Let $f : \mathbb{R} \to [0,\infty)$ be of class C^4. Then $Z^2 = f$ has a C^2 solution if and only if there exists a continuous function γ vanishing on $\{t \in \mathbb{R} : f^{(j)}(t) = 0,\ 0 \le j \le 4\}$ such that for each local minimum t_0 of f with $f(t_0) > 0$ we have $f''(t_0) \le \gamma(t_0)f(t_0)^{1/2}$.*

Bony, Colombini, and Pernazza [19] extended this result. Forcing f and sufficiently many of its derivatives to vanish on all local minima of f, turns out to be sufficient for the existence of C^p square roots:

Theorem 3.25 ([19])**.** *Let $f : \mathbb{R} \to [0,\infty)$ be of class C^{2p}, $p \ge 2$, and assume that f vanishes along with its derivatives up to order $2p-4$ at all its local minima. Then $Z^2 = f$ has a C^p solution. A solution with a derivative of order $p+1$ everywhere exists, provided that, under the same assumption on the minima, f is C^{2p+2}.*

3.3.2. *A regularity class for taking radicals.* A different approach is due to Ray and Schmidt-Hieber [98]. They define a regularity class $\mathcal{F}^\beta$, $\beta > 0$, for nonnegative functions $f : [0,1] \to [0,\infty)$ which behaves nicely with respect to radicals. For $\beta > 0$ we let C^β be a short notation for the Hölder class $C^{m,\alpha}$, where m is the largest integer strictly smaller than β and $\alpha = \beta - m$. For $f \in C^\beta([0,1])$ set

$$|f|_{F^\beta([0,1])} := \max_{1\le j<\beta} \Big(\sup_{t\in[0,1]} \frac{|f^{(j)}(t)|^\beta}{|f(t)|^{\beta-j}} \Big)^{1/j}.$$

It measures the flatness of f near its zeros. In particular, if $|f|_{F^\beta([0,1])} < \infty$ then that f vanishes at some point entails that also all its derivatives up to order $< \beta$ vanish. Consider

$$F^\beta([0,1]) := \{f \in C^\beta([0,1]) : f \ge 0,\ \|f\|_{F^\beta([0,1])} < \infty\},$$

where

$$\|f\|_{F^\beta([0,1])} := \|f\|_{C^\beta([0,1])} + |f|_{F^\beta([0,1])}.$$

Note that $f \in F^\beta([0,1])$ may possess infinitely many nonzero local mimima.

Theorem 3.26 ([98])**.** *Let $\alpha \in (0,1]$ and $\beta > 0$. For all $f \in F^\beta([0,1])$, we have*

$$\|f^\alpha\|_{F^{\alpha\beta}([0,1])} \le C(\alpha,\beta)\, \|f\|^\alpha_{F^\beta([0,1])}.$$

In particular, $f^\alpha \in C^{\alpha\beta}([0,1])$.

Moreover, bounds on the wavelet coefficients of f^α are derived in [98] which give additional information on the local regularity of f^α.

3.3.3. *Finite order of contact.* In the case of a curve of monic hyperbolic polynomials $P_a(t)$, $t \in I$, of degree d, sufficient coefficients for the existence of C^p roots can be given in terms of the differentiability of the coefficients and the finite order of contact of the roots; see Rainer [92, 95].

In order to present these results we need some terminology. The polynomial

$$\Delta_s(X_1,\dots,X_d) := \sum_{i_1<\dots<i_s} (X_{i_1} - X_{i_2})^2 \cdots (X_{i_1} - X_{i_s})^2 \cdots (X_{i_{s-1}} - X_{i_s})^2$$

is symmetric and so it is a (unique) polynomial in the elementary symmetric functions:

$$\Delta_s = \tilde\Delta_s(\sigma_1,\dots,\sigma_d).$$

Thus $t \mapsto \tilde\Delta_s(P_a(t))$, where we identify P_a with its vector of coefficients, is well-defined. We say that $P_a(t)$, $t \in I$, is *normally nonflat* if for each $t_0 \in I$ the following condition is satisfied. Let s be the maximal integer such that the germ at t_0 of $t \mapsto \tilde\Delta_s(P_a(t))$ is not zero. Then

$$m_{t_0}(\tilde\Delta_s(P_a)) := \sup\{m \in \mathbb{N} : (t-t_0)^{-m}\tilde\Delta_s(P_a(t)) \text{ is continuous near } t_0\} < \infty.$$

We call this quantity the *multiplicity of $\tilde\Delta_s(P_a)$ at t_0*; obviously we can analogously define the multiplicity of any continuous univariate function. It is not hard to see that $P_a(t)$, $t \in I$, is normally nonflat if and only if for any continuous system λ_j, $1 \le j \le d$, of its roots $m_{t_0}(\lambda_i - \lambda_j) = \infty$ implies $\lambda_i = \lambda_j$ near t_0.

Let $P_a(t)$, $t \in I$, be a normally nonflat curve of monic hyperbolic polynomials. In [92, 95], numbers $\gamma(P_a), \Gamma(P_a) \in \mathbb{N} \cup \{\infty\}$, $\gamma(P_a) \le \Gamma(P_a)$, are defined, in terms of a splitting algorithm for P_a, that encode the conditions for C^p solvability. We refer the reader to [92] and [95] for the details of the definition.

Theorem 3.27 ([92, 95])**.** *Let $P_a(t)$, $t \in I$, be a normally nonflat curve of monic hyperbolic polynomials of degree d. If $p \in \mathbb{N} \cup \{\infty\}$ and $P_a(t)$, $t \in I$, has $C^{p+\Gamma(P_a)}$ coefficients, then it admits a $C^{p+\gamma(P_a)}$ system of its roots.*

Note that for $p = \infty$ the theorem yields the result of Alekseevsky, Kriegl, Losik, and Michor [2] that C^∞ systems of the roots exist if the coefficients are C^∞.

3.3.4. *Definability: no oscillation.* Let us state one more result. It shows that, actually, it is not the infinite contact between the roots that is to blame for the loss of regularity, but oscillation. Of course, smooth coefficients can admit infinite oscillation only in conjuction with infinite flatness. The result is formulated in the framework of o-minimal expansions $\mathscr{S}$ of the real field. That means that $\mathscr{S}$ is a family $\mathscr{S} = (\mathscr{S}_n)_{n\geq 1}$, where $\mathscr{S}_n$ is a collection of subsets of $\mathbb{R}^n$ such that

- $\mathscr{S}_n$ is a boolean algebra with respect to the usual set-theoretic operations,
- $\mathscr{S}_n$ contains all semialgebraic subsets of $\mathbb{R}^n$,
- $\mathscr{S}$ is stable by cartesian products and linear projections,
- each $S \in \mathscr{S}_1$ has only finitely many connected components.

The *$\mathscr{S}$-definable sets* are those sets S such that there exists $n \geq 1$ with $S \in \mathscr{S}_n$. A map is $\mathscr{S}$-definable if its graph is $\mathscr{S}$-definable.

Theorem 3.28 ([92, Theorem 4.12])**.** *Let $\mathscr{S}$ be an arbitrary o-minimal expansion of the real field. Let $P_a(t)$, $t \in I$, be a curve of monic hyperbolic polynomials with C^∞-coefficients that are $\mathscr{S}$-definable. Then there exists a C^∞-system of the roots of P_a consisting of $\mathscr{S}$-definable functions.*

By Miller's dichotomy theorem [77], for a fixed o-minimal expansion $\mathscr{S}$, either for every $\mathscr{S}$-definable function $f : \mathbb{R} \to \mathbb{R}$ there is $N \in \mathbb{N}$ such that $f(t) = O(t^N)$ as $t \to \infty$, or the global exponential function $\exp : \mathbb{R} \to \mathbb{R}$ is $\mathscr{S}$-definable. In the latter case, there are infinitely flat C^∞ functions in $\mathscr{S}$. On the other hand, $\mathscr{S}$ cannot contain oscillating functions.

3.4. **Gårding hyperbolic polynomials.** There are various ways to generalize univariate hyperbolic polynomials to multivariate polynomials. Of particular interest are the *Gårding hyperbolic polynomials* and the closely related *real stable polynomials*; they have numerous applications in PDEs, combinatorics, optimization, functional analysis, probability, etc.

3.4.1. *Gårding hyperbolic and real stable polynomials.* A homogeneous polynomial $f(Z_1, \dots, Z_n) \in \mathbb{R}[Z_1, \dots, Z_n]$ of degree d is said to be *Gårding hyperbolic with respect to $v \in \mathbb{R}^n$* if $f(v) \neq 0$ and for all $x \in \mathbb{R}^n$ the univariate polynomial $f(x - Tv) \in \mathbb{R}[T]$ is hyperbolic, i.e., has all roots real. Geometrically, this means that any affine line with direction v meets the real hypersurface $\{x \in \mathbb{R}^n : f(x) = 0\}$ in d points (with multiplicities).

This notion was introduced by Gårding [35] in the 1950s. He showed that f being Gårding hyperbolic with respect to a direction v is a necessary and sufficient condition for local well-posedness of the Cauchy problem with principal symbol f and initial data on a hyperplane with normal vector v. Gårding hyperbolic polynomials have found many applications ever since, for instance, in Gurvits' proof [42] of the van der Waerden conjecture.

An important example is the determinant on the real vector space of $d \times d$ Hermitian matrices: it is Gårding hyperbolic with respect to the identity matrix $\mathbb{I}$.

A polynomial $f(Z_1,\dots,Z_n)\in\mathbb{C}[Z_1,\dots,Z_n]$ is called *stable* if $f(z_1,\dots,z_n)\neq 0$ for all $(z_1,\dots,z_n)\in\mathbb{C}^n$ with $\operatorname{Im}(z_j)>0$, $1\le j\le n$. If all coefficients of a stable polynomial are real, it is called *real stable*. Real stable polynomials played a crucial role in the recent proof of the Kadison–Singer conjecture [73].

These notions are closely related.

Proposition 3.29 ([20, Proposition 1.1])**.** *Let $f\in\mathbb{R}[Z_1,\dots,Z_n]$ be of degree d and let $f_H\in\mathbb{R}[Z_1,\dots,Z_n,W]$ be the unique homogeneous polynomial of degree d such that*

$$f_H(Z_1,\dots,Z_n,1)=f(Z_1,\dots,Z_n).$$

Then f is real stable if and only if f_H is Gårding hyperbolic with respect to every vector $v=(v_1,\dots,v_n,0)$ with $v_i>0$ for all i.

Interesting examples of real stable polynomials are generated as follows (see [20, Proposition 1.12]): Let $A_1,\dots,A_n$ be positive semidefinite $d\times d$ matrices and B a (complex) Hermitian $d\times d$ matrix. Then the polynomial

$$f(Z_1,\dots,Z_n)=\det\Big(\sum_{j=1}^n Z_jA_j+B\Big)$$

is either real stable or identically zero. Conversely, any real stable polynomial in two variables Z_1 and Z_2 can be written as $\pm\det(Z_1A_1+Z_2A_2+B)$, where A_j are positive semidefinite and B is symmetric, see [20, Theorem 1.13]. The latter result is based on the Lax conjecture [66] for Gårding hyperbolic polynomials which is true: as noted in [67] it follows from [46] and [113].

Remark 3.30. The Lax conjecture states that a homogeneous polynomial f on $\mathbb{R}^3$ is Gårding hyperbolic of degree d with respect to the direction $(1,0,0)$ with $f(1,0,0)=1$ if and only if there exist real symmetric $d\times d$ matrices A and B such that

$$f(X,Y,Z)=\det(X\mathbb{I}+YA+ZB). \tag{25}$$

A representation of type (25) is in general not possible for Gårding hyperbolic polynomials on $\mathbb{R}^n$ with $n\ge 4$. Indeed, the dimension of the space of Gårding hyperbolic polynomials of degree d on $\mathbb{R}^n$ with respect to a fixed direction is $\binom{n+d-1}{d}$ while the set of polynomials in $\mathbb{R}[X_1,\dots,X_n]$ of the form

$$\det(X_1\mathbb{I}+X_2A_2+\dots+X_nA_n), \tag{26}$$

where A_i are real symmetric $d\times d$ matrices, has dimension at most $(n-1)\cdot\binom{d+1}{2}$. A particular homogeneous polynomial of degree 2 which is Gårding hyperbolic with respect to $(1,0,\dots,0)$ but cannot be represented in the form (26) is the Lorentzian polynomial $f(X)=X_1^2-X_2^2-\dots-X_n^2$ for $n\ge 4$. Cf. [67, p. 2498].

See the survey article [9] for background on Gårding hyperbolic polynomials and more ways to generated examples.

3.4.2. *Characteristic roots.* Let $f(Z_1,\dots,Z_n)\in\mathbb{R}[Z_1,\dots,Z_n]$ be a homogeneous polynomial of degree d which is Gårding hyperbolic with respect to a direction $v\in\mathbb{R}^n$. We may factorize

$$f(x+Tv)=f(v)\prod_{j=1}^d(T+\lambda_j^{\downarrow}(x)),$$

where

$$\lambda_1^\downarrow(x) \geq \ldots \geq \lambda_d^\downarrow(x)$$

are the decreasingly ordered roots of $f(x - Tv) \in \mathbb{R}[T]$. We call

$$\lambda^\downarrow = (\lambda_1^\downarrow, \ldots, \lambda_d^\downarrow) : \mathbb{R}^n \to \mathbb{R}^d$$

the *characteristic map* of f with respect to v and $\lambda_1^\downarrow, \ldots, \lambda_d^\downarrow$ (in no particular order) the *characteristic roots*. The polynomial f is recovered by $f(x) = f(v) \prod_{j=1}^d \lambda_j^\downarrow(x)$. Then, for all $j = 1, \ldots, d$, $r \in \mathbb{R}_{\geq 0}$, and $s \in \mathbb{R}$,

$$\begin{aligned} \lambda_j^\downarrow(rx + sv) &= r\lambda_j^\downarrow(x) + s, \\ \lambda_j^\downarrow(-x) &= -\lambda_{d+1-j}^\downarrow(x). \end{aligned} \tag{27}$$

Gårding [36, Theorem 2] proved that $\lambda_d^\downarrow$ is concave which, by (27), is equivalent to $\lambda_1^\downarrow$ being convex. In view of (27) it follows that the largest root $\lambda_1^\downarrow$ is sublinear (i.e., positively homogeneous and subadditive). The connected component C_f of the set $\mathbb{R}^n \setminus \{f = 0\}$ which contains v is an open convex cone, one has $C_f = \{x \in \mathbb{R}^n : \lambda_d^\downarrow(x) > 0\}$, and f is Gårding hyperbolic with respect to each $w \in C_f$.

Following [9, Theorem 3.1], we can produce new Gårding hyperbolic polynomials from the given f: if g is a homogeneous symmetric polynomial of degree e on $\mathbb{R}^d$ which is Gårding hyperbolic with respect to $(1, 1, \ldots, 1)$ and has characteristic map $\mu^\downarrow$, then $g \circ \lambda^\downarrow$ is Gårding hyperbolic of degree e with respect to v and its characteristic map is $\mu^\downarrow \circ \lambda^\downarrow$.

Let us apply this to the homogeneous symmetric polynomial

$$g_k(Y_1, \ldots, Y_d) := \prod_{\substack{I \subseteq \{1,\ldots,d\} \\ |I|=k}} \sum_{i \in I} Y_i$$

of degree $\binom{d}{k}$ which is Gårding hyperbolic with respect to $(1, 1, \ldots, 1)$ and has the characteristic roots $\mu_I(Y) = \frac{1}{k} \sum_{i \in I} Y_i$, where I ranges over the subsets of $\{1, \ldots, d\}$ with k elements. On the set $\{Y_1 \geq Y_2 \geq \cdots \geq Y_d\}$ we have $\mu_I(Y) \geq \mu_J(Y)$ if and only if $\sum_{i \in I} i \leq \sum_{j \in J} j$, in particular, $\mu_{\{1,\ldots,k\}}$ is the largest characteristic root. Then $g_k \circ \lambda^\downarrow$ is Gårding hyperbolic with respect to v and has the largest characteristic root $\mu_{\{1,\ldots,k\}} \circ \lambda^\downarrow = \frac{1}{k} \sum_{i=1}^k \lambda_i^\downarrow$. From Gårding's result we may conclude that, for all $k = 1, \ldots, d$, the sum of the k largest roots

$$\sigma_k := \sum_{i=1}^k \lambda_i^\downarrow$$

is a sublinear function on $\mathbb{R}^n$. Any finite sublinear function on $\mathbb{R}^n$ is globally Lipschitz; see [101, Corollary 10.5.1]. So the functions σ_k, $k = 1, \ldots, d$, are convex and globally Lipschitz. We have proved

Proposition 3.31 ([96, Proposition 4.1])**.** *Let $f(Z_1, \ldots, Z_n) \in \mathbb{R}[Z_1, \ldots, Z_n]$ be a homogeneous polynomial of degree d which is Gårding hyperbolic with respect to a direction $v \in \mathbb{R}^n$. The characteristic map $\lambda^\downarrow : \mathbb{R}^n \to \mathbb{R}^d$ is globally Lipschitz and difference-convex on $\mathbb{R}^n$.*

A real valued function on a convex subset of $\mathbb{R}^n$ is called *difference-convex* if it can be written as the difference of two continuous convex functions (cf. [7] and [48]). The class DC of difference-convex functions arises as the smallest vector

space containing all the continuous convex function on the given set. If $U \subseteq \mathbb{R}^n$ is open, then (cf. [112, Theorem 11] and [48, Section II])

$$C^{1,1}(U) \subseteq DC_{\mathrm{loc}}(U) \subseteq C^{0,1}(U).$$

On the other hand, the first order partial derivatives of a difference-convex function $f : U \to \mathbb{R}$ have bounded variation, i.e., the weak second order partial derivatives of f are signed Radon measures (cf. Section 7.4). This follows from Dudley's result [33] that a Schwartz distribution is a convex function if and only if its second derivative is a nonnegative matrix valued Radon measure. In dimension one, a real valued function f on a compact interval is difference-convex if and only if f is absolutely continuous and f' has bounded variation. Note also that $DC(\mathbb{R}^n) = DC_{\mathrm{loc}}(\mathbb{R}^n)$, by [44].

Using the regularity properties of difference-convex functions, we may further investigate the regularity of the roots of Gårding hyperbolic polynomials.

Theorem 3.32 ([96, Lemma 4.5 and Theorem 4.6])**.** *Let $f(Z_1, \dots, Z_n) \in \mathbb{R}[Z_1, \dots, Z_n]$ be a homogeneous polynomial of degree d which is Gårding hyperbolic with respect to some direction in $\mathbb{R}^n$. Then:*

(1) *If $x : \mathbb{R} \to \mathbb{R}^n$ is of class C^1, then there exists a differentiable system $\lambda = (\lambda_1, \dots, \lambda_d)$ of the roots of f along x (i.e., $\lambda(t)$ and $\lambda^{\downarrow}(x(t))$ coincide as unordered d-tuples for all t).*

(2) *Let $x \in DC(\mathbb{R}, \mathbb{R}^n) \cap W^{2,1}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^n)$ (for instance, $x \in C^{1,1}(\mathbb{R}, \mathbb{R}^n)$). Then any differentiable system $\lambda = (\lambda_1, \dots, \lambda_d)$ of the roots of f along x is actually of class*

$$\lambda \in C^1(\mathbb{R}, \mathbb{R}^d) \cap DC(\mathbb{R}, \mathbb{R}^d) \cap W^{2,1}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^d).$$

(3) *The result is uniform in the following sense: Let $I \subseteq \mathbb{R}$ be a bounded open interval. Let U be an open neighborhood of the closure of I^{1+k} in $\mathbb{R}^{1+k}$. Suppose that $x : U \to \mathbb{R}^n$ is such that*
   - *x is locally DC on U,*
   - *$x(\cdot, r) \in C^1(\overline{I}, \mathbb{R}^n) \cap W^{2,1}(I, \mathbb{R}^n)$ for all $r \in \overline{I}^k$.*

   Assume that, for each $r \in \overline{I}^k$, a C^1 system $\lambda(\cdot, r)$ of the roots of f along $x(\cdot, r)$ is fixed. Then the family $\lambda(\cdot, r)$, for $r \in \overline{I}^k$, is bounded in $C^1(\overline{I}, \mathbb{R}^d)$ and there is a nonnegative L^1 function $m : I^k \to [0, \infty)$ such that

$$\|\lambda(\cdot, r)\|_{W^{2,1}(I, \mathbb{R}^d)} \le m(r), \qquad \text{for a.e. } r \in I^k.$$

Remark 3.33. It is shown in [45, Theorem 4.2] that for any Gårding hyperbolic polynomial f with respect to $v \in \mathbb{R}^n$ of degree d, $x_0 \in \mathbb{R}^n$, and $w \in C_f$ there is a system $\lambda = (\lambda_1, \dots, \lambda_d) : \mathbb{R} \to \mathbb{R}^d$ of the roots of f along $t \mapsto x_0 + tw$ which is real analytic and such that each $\lambda_j : \mathbb{R} \to \mathbb{R}$ is strictly increasing and surjective. The inverses of the λ_j form a system of the characteristic roots of the Gårding hyperbolic polynomial f with respect to $w \in \mathbb{R}^n$ along $t \mapsto x_0 + tv$. The proof is based on the homogeneity properties (27) and the description of C_f as the set, where $\lambda_1^{\downarrow} \ge \dots \ge \lambda_d^{\downarrow}$ are positive.

3.5. **Eigenvalues of Hermitian matrices.** Let us apply the above findings to the Gårding hyperbolic polynomial det (with respect to the identity matrix $\mathbb{I}$) on the real vector space $\mathrm{Herm}(d)$ of complex $d \times d$ Hermitian matrices. Its characteristic map $\lambda^{\downarrow} = (\lambda_1^{\downarrow}, \dots, \lambda_d^{\downarrow}) : \mathrm{Herm}(d) \to \mathbb{R}^d$ assigns to a Hermitian matrix A its

eigenvalues in decreasing order

$$\lambda_1^{\downarrow}(A) \geq \cdots \geq \lambda_d^{\downarrow}(A).$$

As a special case of Proposition 3.31, we get:

Corollary 3.34 ([96, Corollary 6.1])**.** *The characteristic map $\lambda^{\downarrow} : \mathrm{Herm}(d) \to \mathbb{R}^d$ is globally Lipschitz and difference-convex on $\mathrm{Herm}(d)$. The sum $\sum_{i=1}^k \lambda_i^{\downarrow}$, for $k = 1, \dots, d$, of the k largest eigenvalues is sublinear.*

The following result is a consequence of Theorem 3.32.

Theorem 3.35 ([96, Theorem 6.2])**.** *Let $A : \mathbb{R} \to \mathrm{Herm}(d)$ be a curve of $d \times d$ Hermitian matrices. Then:*

(1) *If A is of class C^1, then there exists a differentiable system $\lambda = (\lambda_1, \dots, \lambda_d)$ of the eigenvalues of A.*
(2) *If A is of class $DC \cap W^{2,1}_{\mathrm{loc}}$ on $\mathbb{R}$ (for instance, if it is of class $C^{1,1}$), then any differentiable system $\lambda = (\lambda_1, \dots, \lambda_d)$ of the eigenvalues of A is actually of class*
$$\lambda \in C^1(\mathbb{R}, \mathbb{R}^d) \cap DC(\mathbb{R}, \mathbb{R}^d) \cap W^{2,1}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^d).$$
(3) *The result in (2) is uniform in the following sense: Let $I \subseteq \mathbb{R}$ be a bounded open interval. Let U be an open neighborhood of the closure of I^{1+k} in $\mathbb{R}^{1+k}$. Suppose that $A : U \to \mathrm{Herm}(d)$ is such that*
   - *A is locally DC on U,*
   - *$A(\cdot, r)$ is of class $C^1 \cap W^{2,1}$ on $\overline{I}$ for all $r \in \overline{I}^k$.*

   Assume that, for each $r \in \overline{I}^k$, a C^1 system $\lambda(\cdot, r)$ of the eigenvalues of $A(\cdot, r)$ is fixed. Then the family $\lambda(\cdot, r)$, for $r \in \overline{I}^k$, is bounded in $C^1(\overline{I}, \mathbb{R}^d)$ and there is a nonnegative L^1 function $m : I^k \to [0, \infty)$ such that
$$\|\lambda(\cdot, r)\|_{W^{2,1}(I, \mathbb{R}^d)} \leq m(r), \qquad \text{for a.e. } r \in I^k.$$

The conclusion of this theorem is best-possible among all Sobolev spaces $W^{k,p}$. Indeed, by the Sobolev inequality, $W^{k,p}$ regularity with $k+p > 3$ would imply $C^{1,\alpha}$ regularity with some $\alpha > 0$, contradicting the following counterexample.

Example 3.36 ([61])**.** There is a C^∞ curve $A(t)$, $t \in \mathbb{R}$, of real symmetric 2×2 matrices and a convergent sequence $t_n \in \mathbb{R}$ such that

$$A(t_n + t) = \begin{pmatrix} 2^{-n^2} & 2^{-n}t \\ 2^{-n}t & -2^{-n^2} \end{pmatrix}, \qquad \text{for } |t| \leq n^{-2},\ n \geq 1.$$

There is no $C^{1,\alpha}$-system of the eigenvalues of A for any $\alpha > 0$.

A little more work leads to the following stronger version of (1).

Theorem 3.37. *If $A : \mathbb{R} \to \mathrm{Herm}(d)$ is of class C^1, then there exists a C^1 system $\lambda = (\lambda_1, \dots, \lambda_d)$ of the eigenvalues of A.*

This is due to Rellich [100] in the case of symmetric matrices (see also Weyl [115]), and it was proved for normal matrices in Rainer [94]; see Theorem 3.43 below.

Note that, conversely, Theorem 3.35 implies Theorem 3.32 if $n \leq 3$; cf. Remark 3.30.

3.6. **Singular values.** Theorem 3.35 contains information on the regularity of singular values. Let $A \in \mathbb{C}^{m\times n}$ be any complex $m \times n$ matrix and let

$$\sigma_1^{\downarrow}(A) \geq \sigma_2^{\downarrow}(A) \geq \cdots \geq \sigma_n^{\downarrow}(A) \geq 0$$

be the singular values of A in decreasing order, i.e., the nonnegative square roots of the eigenvalues of A^*A. If $\operatorname{rank} A = \ell$, then $\sigma_{\ell+1}^{\downarrow}(A) = \cdots = \sigma_n^{\downarrow}(A) = 0$. Thus we set $\ell := \min\{m, n\}$ and consider only $\sigma_j^{\downarrow}(A)$, for $j = 1, \ldots, \ell$. We may consider the $\sigma_j^{\downarrow}$ as functions on the vector space $\mathbb{C}^{m\times n}$.

Without loss of generality assume that $m \leq n$ and let $\tilde{A}$ be the $n \times n$ matrix resulting from A by adding $n - m$ rows consisting of zeros. Then the eigenvalues of the Hermitian matrix

$$\begin{pmatrix} 0 & \tilde{A} \\ \tilde{A}^* & 0 \end{pmatrix} \tag{28}$$

are precisely

$$\sigma_1^{\downarrow}(A) \geq \cdots \geq \sigma_n^{\downarrow}(A) \geq -\sigma_n^{\downarrow}(A) \geq \cdots \geq -\sigma_1^{\downarrow}(A).$$

By Corollary 3.34, for all $k = 1, \ldots, \ell$, the sum

$$\sum_{j=1}^{k} \sigma_j^{\downarrow}$$

of the k largest singular values is a sublinear function on $\mathbb{C}^{m\times n}$ viewed as a real vector space. In fact, the sums $A \mapsto \sum_{j=1}^{k} \sigma_j^{\downarrow}(A)$ are the so-called *Ky Fan norms*.

Corollary 3.38. *The mapping $\sigma^{\downarrow} = (\sigma_1^{\downarrow}, \ldots, \sigma_\ell^{\downarrow}) : \mathbb{C}^{m\times n} \to \mathbb{R}^\ell$, where $\ell = \min\{m, n\}$, is globally Lipschitz and difference-convex on $\mathbb{C}^{m\times n}$.*

Theorem 3.39. *Let $A : \mathbb{R} \to \mathbb{C}^{m\times n}$ be a curve of $m \times n$ complex matrices of class $C^1 \cap DC \cap W^{2,1}_{\mathrm{loc}}$. If $\operatorname{rank} A(t) = \min\{m, n\} =: \ell$ for all t, then there exists a system $\sigma = (\sigma_1, \ldots, \sigma_\ell)$ of the singular values of A such that $\sigma \in C^1(\mathbb{R}, \mathbb{R}^\ell) \cap DC(\mathbb{R}, \mathbb{R}^\ell) \cap W^{2,1}_{\mathrm{loc}}(\mathbb{R}, \mathbb{R}^\ell)$. This result is uniform in the sense explained in Theorem 3.35.*

Proof. Apply Theorem 3.35 to the Hermitian matrix (28). The condition on the rank of A guarantees that the nontrivial singular values of A are always strictly positive and hence there exists a C^1 system of them, since there exists a C^1 system of the eigenvalues of (28). □

The rank condition is necessary; for instance, the singular value of the symmetric 1×1 matrix $A(t) = (t)$ is $|t|$ which does not admit a C^1 parameterization.

3.7. **Eigenvalues of normal matrices.** We have seen that the eigenvalues of Hermitian matrices have similar regularity properties as the roots of hyperbolic polynomials. Actually, they are slightly stronger under much weaker assumptions, see Theorem 3.35, since the eigenvalues of Hermitian matrices are the characteristic roots of the determinant which is Gårding hyperbolic with respect to the identity matrix.

Interestingly, the eigenvalues of the wider class of normal complex matrices have similar strong regularity properties.

Theorem 3.40 ([12], [11, VII.4.1])**.** *Let A, B be normal complex $d \times d$ matrices and let $\lambda_j(A)$ and $\lambda_j(B)$, $1 \le j \le d$, denote the respective eigenvalues. Then*

$$\min_{\sigma \in \mathrm{S}_d} \max_{1 \le j \le d} |\lambda_j(A) - \lambda_{\sigma(j)}(B)| \le C \, \|A - B\|$$

for a universal constant C with $1 < C < 3$, where S_d is the symmetric group and $\|\cdot\|$ is the operator norm.

This result shows that the unordered d-tuple of eigenvalues $[\lambda_1(A), \dots, \lambda_d(A)]$ is Lipschitz continuous as a function of the normal matrix A. On the left-hand side we have the Wasserstein metric W_∞ (up to a constant factor); see Example 6.8. In general, the single eigenvalues do not admit a continuous parameterization, as shown by the following example.

Example 3.41. Each choice of the eigenvalues of the normal matrix

$$A(z) := \begin{pmatrix} 0 & z \\ |z| & 0 \end{pmatrix}, \quad z \in \mathbb{C},$$

must be discontinuous in a neighborhood of 0.

Continuous systems of the eigenvalues always exist for continuous families of Hermitian matrices (e.g., by ordering the eigenvalues increasingly or decreasingly) or for continuous real one-parameter families of normal matrices (cf. [52, II.5.2]). The following theorem should be compared to Bronshtein's theorem 3.2.

Theorem 3.42 ([94, Proposition 6.3])**.** *Let $A(t)$, $t \in \mathbb{R}$, be a $C^{0,1}$ curve of normal complex matrices. Then any continuous system of the eigenvalues of A is $C^{0,1}$.*

Let us sketch the main ideas in the proof. Fix $s \in \mathbb{R}$ and an eigenvalue z of $A(s)$. Let m be the multiplicity of z. For t near s, there are m eigenvalues $\lambda_1(t), \dots, \lambda_m(t)$, close to z and λ_j, $1 \le j \le m$, are continuous functions. For each t, we may choose an orthonormal system of eigenvectors $v_1(t), \dots, v_m(t)$ corresponding to $\lambda_1(t), \dots, \lambda_m(t)$. For each t and each sequence $t_k \to t$, we have, after passing to a subsequence, $v_j(t_k) \to w_j(t)$, where $w_1(t), \dots, w_m(t)$ form an orthonormal system of eigenvectors of $A(t)$ (restricted to the direct sum of eigenspaces of the $\lambda_j(t)$, $1 \le j \le m$). Then

$$\frac{A(t) - \lambda_j(t)}{t_k - t} v_j(t_k) + \frac{A(t_k) - A(t)}{t_k - t} v_j(t_k) = \frac{\lambda_j(t_k) - \lambda_j(t)}{t_k - t} v_j(t_k).$$

By taking the inner product with $w_j(t)$ and passing to the limit, one can show that for any continuous system λ_j of the eigenvalues of A the derivatives $\lambda_j'(t)$ exist whenever $A'(t)$ exists, they form the set of eigenvalues of $A'(t)$ and satisfy

$$\lambda_j'(t) = \langle A'(t) w_j(t) \mid w_j(t) \rangle, \tag{29}$$

possibly after applying a suitable permutation of the eigenvalues on one side of t. Now it suffices to check that λ_j is locally absolutely continuous. Then, by means of (29), we see that λ_j' exists almost everywhere and is locally bounded. We may conclude that λ_j is $C^{0,1}$.

With the help of the formula (29), one can show that there are C^1 systems of the eigenvalues provided that $t \mapsto A(t)$ is C^1.

Theorem 3.43 ([94, Proposition 6.11])**.** *Let $A(t)$, $t \in \mathbb{R}$, be a C^1 curve of normal complex matrices. Then there is a C^1 system of the eigenvalues of A. If $t \mapsto A(t)$*

is C^2, then there is a C^1 system of the eigenvalues of A which is twice differentiable everywhere.

Note that this is best possible, by Example 3.36. The next example shows that the assumption that A is normal cannot be omitted.

Example 3.44 ([52, II.5.9])**.** Let $\alpha > 1$ and $\beta > 2$. Then

$$A(t) = \begin{pmatrix} |t|^\alpha & |t|^\alpha - |t|^\beta\big(2 + \sin\frac{1}{|t|}\big) \\ -|t|^\alpha & -|t|^\alpha \end{pmatrix}, \quad t \in \mathbb{R} \setminus \{0\}, \quad A(0) = 0,$$

forms a C^1 curve of diagonalizable (but not normal) matrices. The eigenvalues of A are given by

$$\lambda_\pm(t) = \pm|t|^{\frac{\alpha+\beta}{2}}\Big(2 + \sin\frac{1}{|t|}\Big)^{\frac{1}{2}}, \quad t \in \mathbb{R} \setminus \{0\}, \quad \lambda_\pm(0) = 0.$$

The derivatives $\lambda'_\pm$ exist everywhere, but they are discontinuous at 0 if $\alpha + \beta \le 4$ and even unbounded near 0 if $\alpha + \beta < 4$.

For multiparameter families of normal matrices we have the following.

Corollary 3.45 ([94, Theorem 6.19])**.** *Let $A(x)$, $x \in U$, be a $C^{0,1}$ family of normal complex matrices on an open subset $U \subseteq \mathbb{R}^n$ and let $\lambda : V \to \mathbb{C}$ be a continuous eigenvalue of A defined on an open subset $V \subseteq U$. Then λ is of class $C^{0,1}$. Moreover, if $x_0 \in U \cap \overline{V}$ and $c : \mathbb{R} \to U$ is a $C^{0,1}$ curve with $c(0) = x_0$ and $c((0,1)) \subseteq V$, then $\lambda \circ c|_{(0,1)}$ is globally Lipschitz on $(0,1)$.*

The first statement follows from Theorem 3.42 and the fact that $C^{0,1}$ regularity can be tested on C^∞ curves (cf. [60]). For the second assertion, we note that the formula (29) yields that the Lipschitz constant of $\lambda \circ c|_{(0,1)}$ is globally bounded by the one of $A \circ c|_{(0,1)}$.

In the spirit of Theorem 3.27, it is possible to give sufficient conditions for the existence of C^p eigenvalues and eigenvectors of curves of normal matrices in terms of the differentiability of the entries and the order of contact of the eigenvalues.

Theorem 3.46 ([95, Theorem 7.8])**.** *Let $A(t)$, $t \in I$, be a curve of normal complex matrices. Suppose that the characteristic polynomial is normally nonflat. There exists $\Theta(A) \in \mathbb{N} \cup \{\infty\}$ such that if $p \in \mathbb{N}_{\ge 1} \cup \{\infty\}$ and $A(t)$, $t \in I$, has $C^{p+\Theta(A)}$ entries, then it admits a $C^{p+\Theta(A)}$ system of its eigenvalues and a C^p system of eigenvectors.*

If the entries of the normal matrix A are definable in an o-minimal expansion of the real field, then the eigenvalues have the same regularity without the assumption of normal nonflatness.

Theorem 3.47 ([95, Theorem 8.1])**.** *Let $A(t)$, $t \in I$, be a definable curve of normal complex matrices. There exists an integer $\Theta(A) \in \mathbb{N}$ such that if $p \in \mathbb{N} \cup \{\infty\}$ and $A(t)$, $t \in I$, has $C^{p+\Theta(A)}$ entries, then it admits a $C^{p+\Theta(A)}$ system of its eigenvalues.*

That there is no hope for continuous eigenvectors in this setup (even if there is no oscillation) is seen by the following example.

Example 3.48 ([95, Example 9.5])**.** Let

$$B(t) := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{for } t \ge 0 \quad \text{and} \quad B(t) := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{for } t < 0.$$

Then $A(t) := t^{p+1}B(t)$ is a definable C^p curve of real symmetric matrices which does not admit a continuous system of eigenvectors. If the underlying o-minimal expansion of the real field defines the exponential function, then $A(t) := e^{-1/t^2}B(t)$ is a definable C^∞ curve of real symmetric matrices not admitting continuous eigenvectors.

Remark 3.49. Many of the results are true (in slightly adjusted form) for parameterized families of unbounded normal operators in Hilbert space with common domain of definition and with compact resolvent. Cf. [62, 63, 94] and references therein.

4. Regularity of the roots in the general (nonhyperbolic) case

Let $I \subseteq \mathbb{R}$ be a bounded open interval and let

$$P_a(t)(Z) = Z^d + \sum_{j=1}^{d} a_j(t) Z^{d-j}, \quad t \in I, \tag{30}$$

be a monic polynomial with complex valued functions $a_j : I \to \mathbb{C}$, $j = 1, \dots, d$, as coefficients. It is not hard to see that if the coefficients a_j are continuous then P_a always admits a continuous system of the roots (e.g. [52, II.5.2]). In the absence of hyperbolicity assumptions, the roots of P cannot be chosen Lipschitz even in the simplest radical case: $a_d = t$ and $a_j \equiv 0$ for $j \le d-1$.

Motivated by the analysis of certain systems of pseudo-differential equations, Spagnolo [106] posed the question of absolute continuity of the roots of P_a. Using explicit formulas for the roots, he gave the positive answer for the polynomials of degree 2 and 3 in [105].

The positive answer to this question in the general case was given in Parusiński and Rainer [83, 84] with two different proofs. They both use in a substantial way the work of Ghisi and Gobbino [37] on the radical case, where the optimal regularity conditions were given. The paper [84] also contains the optimal Sobolev regularity of the roots in the general case, see below for precise statements.

4.1. **The case of radicals.** The first result towards absolute continuity of the roots is probably Lemma 1 in Colombini, Jannelli, and Spagnolo [25] which states that for a real valued nonnegative function f of class $C^{k,\alpha}(\overline{I})$ on a bounded open interval I, with $k \in \mathbb{N}_{\ge 1}$ and $0 < \alpha \le 1$, the real radical $f^{1/(k+\alpha)}$ is absolutely continuous on I and satisfies

$$\|(f^{\frac{1}{k+\alpha}})'\|_{L^1(I)}^{k+\alpha} \le C(k,\alpha,I)\|f\|_{C^{k,\alpha}(\overline{I})}.$$

Tarama [108] extended this lemma to real valued functions (not necessarily nonnegative). A better summability for the weak partial derivatives of $f^{1/(k+1)}$ was obtained by Colombini and Lerner [26] for nonnegative C^{k+1} functions f of several real variables.

The case of radicals was settled by Ghisi and Gobbino [37] by finding their optimal regularity.

Theorem 4.1 (Ghisi and Gobbino [37, Theorem 2.2])**.** *Let $I \subseteq \mathbb{R}$ be a bounded open interval and let f be a real valued continuous function such that there exists $g \in C^{k,\alpha}(\overline{I})$ so that $|f|^{k+\alpha} = |g|$ on I. Then $f' \in L^p_w(I)$, where $\frac{1}{p} + \frac{1}{k+\alpha} = 1$, and*

$$\|f'\|_{p,w,I} \le C(k) \max\left\{ [\text{Höld}_{\alpha,I}(g^{(k)})]^{\frac{1}{k+\alpha}} |I|^{\frac{1}{p}}, \|g'\|_{L^\infty(I)}^{\frac{1}{k+\alpha}} \right\}; \tag{31}$$

in particular, f belongs to the Sobolev space $W^{1,q}(I)$ for every $q \in [1,p)$.

Here $L^p_w(I)$ denotes the weak Lebesgue space equipped with the quasinorm $\|f\|_{p,w,I} := \sup_{r\geq 0}\left\{r\cdot\mathcal{L}^1(\{t\in I : |f(t)|>r\})^{\frac{1}{p}}\right\}$, where $\mathcal{L}^1$ is the one dimensional Lebesgue measure. By $\text{Höld}_{\alpha,I}(g^{(k)})$ we mean the α-Hölder constant of $g^{(k)}$ on I, and $|I| = \mathcal{L}^1(I)$ is the length of the interval I, see also Section 7.

Ghisi and Gobbino also provided examples that show that the assumptions as well as the conclusion in their theorem are best possible.

- Take $g(t) = t$ and $f(t) = |t|^{\frac{1}{k+\alpha}}$ on the interval $I = (-1,1)$. Then $f' \notin L^p(I)$ for p defined by $\frac{1}{p}+\frac{1}{k+\alpha} = 1$. So the summability of f' in Theorem 4.1 is optimal. See [37, Example 4.3].
- In [37, Example 4.4], a C^∞ function $g : I \to \mathbb{R}$ on an interval $I = (0,r)$ with the following properties is constructed:
  - g belongs to $C^{k,\beta}(\overline{I})$ for every $\beta < \alpha$ but not for $\beta = \alpha$.
  - $f := |g|^{\frac{1}{k+\alpha}}$ has unbounded variation on I and hence $f' \notin L^1(I)$.

  So the differentiability assumption on g in Theorem 4.1 is optimal.

Higher order Glaeser inequalities [37, Theorems 2.1 and 3.2] play an important role in the proof of Theorem 4.1. We recall a slightly more general statement in the next subsection, see Corollary 4.4.

4.2. **Optimal Sobolev regularity of the roots.**

Theorem 4.2 ([84, Theorem 1])**.** *Let $I \subseteq \mathbb{R}$ be a bounded open interval and let P_a be a monic polynomial (30) with complex valued coefficients $a_j \in C^{d-1,1}(\overline{I})$, $j = 1,\dots,d$. Let $\lambda \in C^0(I)$ be a continuous root of P_a. Then λ is absolutely continuous on I and belongs to the Sobolev space $W^{1,p}(I)$ for every $1 \leq p < d/(d-1)$. The derivative λ' satisfies*

$$\|\lambda'\|_{L^p(I)} \leq C(d,p)\max\{1,|I|^{1/p}\}\max_{1\leq j\leq d}\|a_j\|^{1/j}_{C^{d-1,1}(\overline{I})}, \tag{32}$$

where the constant $C(d,p)$ depends only on d and p.

The above result can be understood as a complex analogue of Bronshtein's theorem, see Theorem 3.2. Its proof given in [84] follows the strategy of the proof of Bronshtein's theorem as explained in Section 3.2, though many details are significantly more complex. The universal splitting $P_{\tilde{a}} = P_b P_c$ allows the induction on the degree of P_a. After the Tschirnhausen transformation $P_b \rightsquigarrow P_{\tilde{b}}$ the new polynomials $P_{\tilde{b}}$ split again near points $t_1 \in I$, where not all $\tilde{b}_i$ vanish. The central idea of the underlying induction is to show an inequality between the combined coefficients of $P_{\tilde{a}}$ and $P_{\tilde{b}}$ and a similar expression for $P_{\tilde{b}}$ and the polynomial issued from the splitting of $P_{\tilde{b}}$, that is [84, Formula (18)]. The proof of this technical step is based on the radical case, Theorem 4.1, and on the following interpolation lemma.

Lemma 4.3 ([84, Lemma 4])**.** *Let $I \subseteq \mathbb{R}$ be a bounded open interval, $m \in \mathbb{N}_{>0}$, and $\alpha \in (0,1]$. If $f \in C^{m,\alpha}(\overline{I})$, then for all $t \in I$ and $s = 1,\dots,m$,*

$$|f^{(s)}(t)| \leq C|I|^{-s}\big(V_I(f) + V_I(f)^{(m+\alpha-s)/(m+\alpha)}(\text{Höld}_{\alpha,I}(f^{(m)}))^{s/(m+\alpha)}|I|^s\big), \tag{33}$$

for a universal constant C depending only on m and α.

Here $V_I(f) := \sup_{t,s\in I}|f(t)-f(s)|$. As a corollary we get the following generalizations of the higher order Glaeser inequalities of [37].

Corollary 4.4 ([84, Corollary 3])**.** *Let $m \in \mathbb{N}_{>0}$ and $\alpha \in (0,1]$. Let $I = (t_0 - \delta, t_0 + \delta)$ with $t_0 \in \mathbb{R}$ and $\delta > 0$. If $f \in C^{m,\alpha}(\bar{I})$ is such that f and f' do not change their sign on I, then for all $s = 1, \dots, m$,*

(34)
$$|f^{(s)}(t_0)| \le C|I|^{-s}\big(|f(t_0)| + |f(t_0)|^{(m+\alpha-s)/(m+\alpha)}(\mathrm{Höld}_{\alpha,I}(f^{(m)}))^{s/(m+\alpha)}|I|^s\big),$$

for a universal constant C depending only on m and α.

Note that the bound (32) is not invariant under rescaling of the interval I, in contrast to (31).

Open Problem 4.5. *Replace (32) by a scale invariant bound.*

Unlike in the case of radicals, i.e., Theorem 4.1, it is not known whether in the setting of Theorem 4.2, λ' belongs to the weak Lebesgue space $L^{d/(d-1)}_w(I)$. The proof cannot be, at least easily, adapted because $\|\cdot\|^p_{p,w,I}$ is not σ-additive.

Open Problem 4.6. *Is λ' in the setting of Theorem 4.2 an element of the weak Lebesgue space $L^{d/(d-1)}_w(I)$? If so is there an explicit bound for $\|\lambda'\|_{d/(d-1),w,I}$ in terms of the coefficients a_j and the interval I?*

4.2.1. *Multivalued Sobolev functions.* In [4] Almgren developed a theory of d-valued Sobolev functions. A d-valued function is a mapping with values in the set $\mathcal{A}_d(\mathbb{R}^m)$ of unordered d-tuples of points in $\mathbb{R}^m$. Let us denote by $[x] = [x_1, \dots, x_d]$ the unordered d-tuple consisting of $x_1, \dots, x_d \in \mathbb{R}^m$; then $[x_1, \dots, x_d] = [x_{\sigma(1)}, \dots, x_{\sigma(d)}]$ for each permutation $\sigma \in \mathrm{S}_d$. The set $\mathcal{A}_d(\mathbb{R}^m) = \{[x] = [x_1, \dots, x_d] : x_i \in \mathbb{R}^m\}$ forms a complete metric space when endowed with the metric

$$d([x],[y]) := \min_{\sigma \in \mathrm{S}_d} \Big(\sum_{i=1}^d |x_i - y_{\sigma(i)}|^2\Big)^{1/2}.$$

This is (up to a constant factor) the Wasserstein metric W_2, see Example 6.8. Almgren proved that there is an integer $N = N(d,m)$, a positive constant $C = C(d,m)$, and an injective Lipschitz mapping $\Delta : \mathcal{A}_d(\mathbb{R}^m) \to \mathbb{R}^N$, with Lipschitz constant $\mathrm{Lip}(\Delta) \le 1$ and $\mathrm{Lip}(\Delta|^{-1}_{\Delta(\mathcal{A}_d(\mathbb{R}^m))}) \le C$; moreover, there is a Lipschitz retraction of $\mathbb{R}^N$ onto $\Delta(\mathcal{A}_d(\mathbb{R}^m))$.

One can use this bi-Lipschitz embedding to define Sobolev spaces of d-valued functions: for open $U \subseteq \mathbb{R}^n$ and $1 \le p \le \infty$ define

$$W^{1,p}(U, \mathcal{A}_d(\mathbb{R}^m)) := \{f : U \to \mathcal{A}_d(\mathbb{R}^m) : \Delta \circ f \in W^{1,p}(U, \mathbb{R}^N)\}.$$

For an intrinsic definition, see [30, Definition 0.5 and Theorem 2.4].

Let us identify $\mathbb{R}^2 \cong \mathbb{C}$. Theorem 4.2 implies a sufficient condition for a d-valued function $U \to \mathcal{A}_d(\mathbb{C})$ to belong to the Sobolev spaces $W^{1,p}(U, \mathcal{A}_d(\mathbb{C}))$ for every $1 \le p < d/(d-1)$; see Theorem 4.7 below.

We shall use the following terminology. By a *parameterization* of a d-valued function $f : U \to \mathcal{A}_d(\mathbb{C})$ we mean a function $\varphi : U \to \mathbb{C}^d$ such that $f(x) = [\varphi(x)] = [\varphi_1(x), \dots, \varphi_d(x)]$ for all $x \in U$. Let $\pi : \mathbb{C}^d \to \mathcal{A}_d(\mathbb{C})$ be defined by $\pi(z) := [z]$; it is a Lipschitz mapping with Lipschitz constant $\mathrm{Lip}(\pi) = 1$. Then a parameterization of f amounts to a lifting φ of f over π, i.e., $f = \pi \circ \varphi$. The elementary symmetric polynomials induce a bijective mapping $a : \mathcal{A}_d(\mathbb{C}) \to \mathbb{C}^d$,

$$a_j([z_1, \dots, z_d]) := (-1)^j \sum_{i_1 < \dots < i_j} z_{i_1} \cdots z_{i_j}, \quad 1 \le j \le d.$$

In other words, the monic complex polynomials of degree d are in one-to-one correspondence with their unordered d-tuples of roots.

Theorem 4.7 ([84, Theorem 6])**.** *Let $U \subseteq \mathbb{R}^n$ be open and let $f : U \to \mathcal{A}_d(\mathbb{C})$ be continuous. If $a \circ f \in C^{d-1,1}(U, \mathbb{C}^d)$, then $f \in W^{1,p}(V, \mathcal{A}_d(\mathbb{C}))$ for each relatively compact open $V \Subset U$ and each $1 \le p < d/(d-1)$. Moreover,*

$$\|\nabla(\Delta \circ f)\|_{L^p(V)} \le C(d, n, p, \mathcal{K}, \Delta)\big(1 + \max_{1\le j\le d} \|a_j \circ f\|^{1/j}_{C^{d-1,1}(\overline{W})}\big),$$

where $\mathcal{K}$ is any finite cover of $\overline{V}$ by open boxes $\prod_{i=1}^d (\alpha_i, \beta_i)$ contained in U and $W = \bigcup \mathcal{K}$.

In particular, the roots of a polynomial P_a of degree d with coefficients $a_j \in C^{d-1,1}(U)$, $j = 1, \dots, d$, form a d-valued function $\lambda : U \to \mathcal{A}_d(\mathbb{C})$ which belongs to $W^{1,p}_{\mathrm{loc}}(U, \mathcal{A}_d(\mathbb{C}))$ for each $1 \le p < d/(d-1)$; in fact, it is well-known that $\lambda : U \to \mathcal{A}_d(\mathbb{C})$ is continuous (cf. [52] or [91, Theorem 1.3.1]). Theorem 4.7 implies that the push-forward

$$(a^{-1})_* : C^{d-1,1}(U, \mathbb{C}^d) \to \bigcap_{1\le p<d/(d-1)} W^{1,p}_{\mathrm{loc}}(U, \mathcal{A}_d(\mathbb{C})).$$

is a bounded mapping. Similarly to the hyperbolic case, the continuity of this map is an important open problem, compare to Example 3.6, Remark 3.7, Open Problem 3.8.

Open Problem 4.8. *Is the map $(a^{-1})_*$ continuous?*

4.3. **Absolute continuity via desingularization. Formulas for the roots.** The main theorem of Parusiński and Rainer [83] also implies the absolute continuity of the roots of (30) provided the coefficients a_j are of class C^k, for k sufficiently large. Its statement, see Theorem 4.9 below, is weaker than the optimal bound of Theorem 4.2. On the other hand, it presents a different approach through modifications of the space of coefficients of the universal monic polynomial that replaces the arguments for each curve of polynomials separately. This approach can be better suited in some cases, for instance in the proof of the bounded variation property, see Theorem 4.11 below.

Theorem 4.9 ([83, Main Theorem])**.** *For every $d \in \mathbb{N}_{>0}$ there is $k = k(d) \in \mathbb{N}_{>0}$ and $p = p(d) > 1$ such that the following holds. Let $I \subseteq \mathbb{R}$ be a bounded open interval and let P_a be a monic polynomial of degree d with complex valued coefficients $a_j \in C^k(\overline{I})$, $j = 1, \dots, d$. Then:*

(1) *Let $\lambda_j \in C^0(I)$, $j = 1, \dots, d$, be a continuous parameterization of the roots of P_a on I. Then the distributional derivative of each λ_j in I is a measurable function $\lambda_j' \in L^q(I)$ for every $q \in [1, p)$. In particular, each $\lambda_j \in W^{1,q}(I)$ for every $q \in [1, p)$.*

(2) *This regularity of the roots is uniform. Let $\{P_{a_\nu} : \nu \in \mathcal{N}\}$,*

$$P_{a_\nu(t)}(Z) = Z^d + \sum_{j=1}^d a_{\nu,j}(t) Z^{d-j} \in C^k(\overline{I})[Z],\ \nu \in \mathcal{N},$$

be a family of curves of polynomials, indexed by ν in some set $\mathcal{N}$, so that the set of coefficients $\{a_{\nu,j} : \nu \in \mathcal{N}, j = 1, \dots, d\}$ is bounded in $C^k(\overline{I})$.

Then the set

$$\{\lambda_\nu \in C^0(I) : P_{a_\nu}(\lambda_\nu) = 0 \text{ on } I,\ \nu \in \mathcal{N}\}$$

is bounded in $W^{1,q}(I)$ *for every* $q \in [1, p)$.

For instance, for $d = 3$ the proof given in [83, Part 3] provides $k(3) = 6$ and $p(3) = 6/5$, which is worse than the optimal $k = 3$, $p = 3/2$, but still better than Spagnolo's result which gives the absolute continuity (i.e. $p = 1$) for $k = 25$.

The proof of the above theorem is based on formulas for the roots of the universal monic polynomial $P_a(Z) = Z^d + \sum_{j=1}^d a_j Z^{d-j}$ in terms of its coefficients $a = (a_1, \dots, a_d) \in \mathbb{C}^d$. They are obtained as follows. By Hironaka's resolution of singularities [49], there is a tower of smooth principalizations

$$\mathbb{C}^d \xleftarrow{\sigma_2} M_2 \xleftarrow{\sigma_{3,2}} M_3 \xleftarrow{\sigma_{4,3}} \cdots \xleftarrow{\sigma_{d,d-1}} M_d \tag{35}$$

which successively principalize the generalized discriminant ideals $\mathscr{D}_m$ in $\mathbb{C}[a]$, $m = 2, \dots, d$, defined in [83, Section 1.1]. These ideals give the stratification of the space of polynomials by root multiplicities. Then, locally on M_d, the roots of the pulled back polynomial $P_{\sigma_d^*(a)}$ are given by linear combinations with rational coefficients

$$\sum_{m=1}^d A_m\,(\varphi_m \circ \sigma_{d,m}), \qquad \text{where} \tag{36}$$
$$\varphi_m = f_m^{\alpha_m} \psi_m(y_{m,1}^{1/q_m}, \dots, y_{m,r_m}^{1/q_m}, y_{m,r_m+1}, \dots, y_{m,d}).$$

Here $\sigma_m = \sigma_2 \circ \sigma_{3,2} \circ \cdots \circ \sigma_{m,m-1}$, $\sigma_{d,m} = \sigma_{m+1,m} \circ \cdots \circ \sigma_{d,d-1}$, $f_m \in \mathscr{D}_m$ is a local generator of $\sigma_m^*(\mathscr{D}_m)$, ψ_m is a convergent power series, $q_m \in \mathbb{N}_{\geq 1}$, $\alpha_m \in \frac{1}{q_m}\mathbb{N}_{\geq 1}$, and $y_{m,i}$ is a local system of coordinates so that $f_m^{-1}(0)$ is given by $y_{m,1} y_{m,2} \cdots y_{m,r_m} = 0$. These formulas reduce the problem to radicals and then one may use Theorem 4.1.

The theorem is first proven for, sufficiently generic, real analytic curves of polynomials. Let $a(t)$ be such a curve, for genericity it suffices that the discriminant of $P_{a(t)}$ is not identically equal to zero. Note that $a(t)$ can be uniquely lifted over the sequence of blowings-up (35) by the universal property of blowing-ups. The main difficulty is to get uniform bounds for the norms independent of the choice of the curve.

If $a(t)$ is arbitrary, then by the Weierstrass approximation theorem it can be approximated by a sequence of polynomial curves $(a_\nu) \subseteq C^\omega(\bar{I}, \mathbb{C}^d)$, such that

$$a_\nu \to a \quad \text{in } C^k(\bar{I}, \mathbb{C}^d) \quad \text{as } \nu \to \infty.$$

Then, thanks to the uniformity of the bounds, one concludes by the Arzelá–Ascoli theorem, or alternatively by the Rellich–Kondrachov compactness theorem, by passing to a subsequence.

4.4. **Multiparameter case.** Though Theorems 4.1, 4.2, 4.9 are one dimensional, they can be fairly easily extended to the multiparameter case by a standard sectioning argument. We recall the statement for the analogue of Theorem 4.2. For Theorem 4.1 and Theorem 4.9, see [37, Theorem 2.3] and [83, Multiparameter Theorem], respectively.

Theorem 4.10 ([84, Theorem 2])**.** *Let $U \subseteq \mathbb{R}^n$ be open and let*

$$P_a(x)(Z) = Z^d + \sum_{j=1}^{d} a_j(x) Z^{d-j}, \quad x \in U, \tag{37}$$

be a monic polynomial with complex valued coefficients $a_j \in C^{d-1,1}(U)$, $j = 1, \dots, d$. Let $\lambda \in C^0(V)$ be a root of P_a on a relatively compact open subset $V \Subset U$. Then λ belongs to the Sobolev space $W^{1,p}(V)$ for every $1 \le p < d/(d-1)$. The distributional gradient $\nabla\lambda$ satisfies

$$\|\nabla\lambda\|_{L^p(V)} \le C(d,n,p,\mathcal{K}) \max_{1\le j\le d} \|a_j\|^{1/j}_{C^{d-1,1}(\overline{W})}, \tag{38}$$

where $\mathcal{K}$ is any finite cover of $\overline{V}$ by open boxes $\prod_{i=1}^n (\alpha_i, \beta_i)$ contained in U and $W = \bigcup \mathcal{K}$; the constant $C(d,n,p,\mathcal{K})$ depends only on d, n, p, and the cover $\mathcal{K}$.

There is a simpler version of the above theorem, avoiding covers by open boxes in the statement, valid for bounded Lipschitz domains, see [85, Theorem A.1].

4.4.1. *Bounded variation of the roots.* There is nevertheless an important difference between the one parameter and the multiparameter case. Recall that by a theorem of Kato [52, II.5.2] a continuous family of complex monic polynomials depending on one real parameter always admits a continuous choice of roots. This is no longer true if the dimension of the parameter space is at least two. In that case, monodromy may prevent the existence of continuous roots. For instance, the radical $\mathbb{R}^2 = \mathbb{C} \ni (x+iy) \mapsto (x+iy)^{1/d}$, $d > 1$, does not admit continuous parameterizations near the origin.

Functions of bounded variation (BV) are integrable functions whose distributional derivative is a vector valued finite Radon measure (cf. Section 7.4). They form an algebra of discontinuous functions. Due to their ability to deal with discontinuities, they are widely used in the applied sciences, see e.g. [53]. It is shown by Parusiński and Rainer in [85] that the roots of a polynomial (37) with coefficients in a differentiability class of sufficiently high order can be represented by functions of bounded variation.

Theorem 4.11 ([85, Theorem 1.1])**.** *For all integers $d, n \ge 2$ there exists an integer $k = k(d,n) \ge \max\{d,n\}$ such that the following holds. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded Lipschitz domain and let (37) be a monic polynomial of degree d with complex valued coefficients $a = (a_1, \dots, a_d) \in C^{k-1,1}(\overline{\Omega}, \mathbb{C}^d)$.*

Then the roots of P_a admit a parameterization $\lambda = (\lambda_1, \dots, \lambda_d)$ by special functions of bounded variation (SBV) on Ω such that

$$\|\lambda\|_{BV(\Omega)} \le C(d,n,\Omega) \max\left\{1, \|a\|_{L^\infty(\Omega)}\right\} \max\left\{1, \|a\|_{C^{k-1,1}(\overline{\Omega})}\right\}. \tag{39}$$

There is a finite collection of C^{k-1} hypersurfaces E_j in Ω such that λ is continuous in the complement of $E := \bigcup_j E_j$. Any hypersurface E_j is closed in an open subset of Ω but possibly not in Ω itself. All discontinuities of λ are jump discontinuities.

The main difficulty of the problem is to make a good choice of the discontinuity set E of the roots. On the complement of E, the roots are of optimal Sobolev class $W^{1,p}$, for all $1 \le p < \frac{d}{d-1}$, by [84], see Theorem 4.10. In general, it is not possible to choose the discontinuity set E of finite codimension one Hausdorff measure $\mathcal{H}^{n-1}$, [85, Example 2.4]. Thus, in order to have bounded variation it is crucial that the

jump height of a selection of a root is integrable (with respect to $\mathcal{H}^{n-1}$) along its discontinuity set, see also the end of Section 6.1.

The proof is based on the radical case, see [85, Theorem 1.4], which is significantly simpler, and on the formulas (36) for the roots of the universal polynomial P_a, $a \in \mathbb{C}^d$, obtained in [83]. The result of Ghisi and Gobbino [37] is used in several places. Interestingly, the method of [83] seems to be better suited for the control of the discontinuities and integrability along them than a more elementary method of [84].

5. Lifting maps over invariants of group representations

5.1. **A reformulation of the regularity problem for hyperbolic polynomials.** The regularity problem for the roots of hyperbolic polynomials can be seen as a special lifting problem: let the symmetric group S_d act on $\mathbb{R}^d$ by permuting the coordinates. The algebra of invariant polynomials $\mathbb{R}[\mathbb{R}^d]^{\mathrm{S}_d}$ is generated by the elementary symmetric functions

$$\sigma_i = \sum_{j_1<\cdots<j_i} x_{j_1}\cdots x_{j_i}, \quad 1 \le i \le d.$$

Each point in the image of the map $\sigma = (\sigma_1, \dots, \sigma_d) : \mathbb{R}^d \to \mathbb{R}^d$ represents, in view of Vieta's formulas, a monic hyperbolic polynomial P_a of degree d, in a unique way. The fiber of σ over this point is the orbit $\mathrm{S}_d(\lambda_1, \dots, \lambda_d) = \{(\lambda_{\tau(1)}, \dots, \lambda_{\tau(d)}) : \tau \in \mathrm{S}_d\}$, where $\lambda_1, \dots, \lambda_d$ are the roots of P_a (with multiplicities). Thus $\sigma(\mathbb{R}^d) = \mathrm{Hyp}(d)$.

In this picture, a family of monic hyperbolic polynomials $P_a(x)$, $x \in U \subseteq \mathbb{R}^n$, is a map $p : U \to \mathbb{R}^d$ with image contained in $\sigma(\mathbb{R}^d)$ and a system of roots for the family P_a is a lifting $\overline{p} : U \to \mathbb{R}^d$ of p over σ, i.e., $\sigma \circ \overline{p} = p$. The regularity problem can be rephrased as follows: Given that p has some regularity, how regular can a lifting of p be chosen?

$$\begin{tikzcd} & \mathbb{R}^d \arrow[d, "\sigma"] \\ U \arrow[ur, dashed, "\overline{p}"] \arrow[r, "p"'] & \sigma(\mathbb{R}^d) \end{tikzcd}$$

This setup can be generalized considerably and we will see that Bronshtein's theorem 3.2 and many other results on the perturbation theory of hyperbolic polynomials hold in greater generality.

We shall also discuss a generalization of the nonhyperpolic case along the same lines in Section 5.5 and subsequent sections.

5.2. **Orthogonal representations of compact Lie groups.** Let G be a compact Lie group and let $\rho : G \to \mathrm{O}(V)$ be an orthogonal representation in a real finite dimensional Euclidean vector space V with inner product $\langle \cdot \mid \cdot \rangle$. We will often just write $G \circlearrowleft V$. In particular, G may be a finite group. By a classical theorem of Hilbert and Nagata, the algebra $\mathbb{R}[V]^G$ of invariant polynomials on V is finitely generated. So let $\{\sigma_i\}_{i=1}^n$ be a system of homogeneous generators of $\mathbb{R}[V]^G$ and call it a system of *basic invariants*.

A system of basic invariants $\{\sigma_i\}_{i=1}^n$ is called *minimal* if there is no polynomial relation of the form $\sigma_i = P(\sigma_1, \dots, \widehat{\sigma_i}, \dots, \sigma_n)$, or equivalently, $\{\sigma_i\}_{i=1}^n$ induces a basis of the real vector space $\mathbb{R}[V]^G_+/(\mathbb{R}[V]^G_+)^2$, where $\mathbb{R}[V]^G_+ = \{f \in \mathbb{R}[V]^G : f(0) =$

0}; cf. [31, Section 3.6]. The elements in a minimal system of basic invariants may not be unique, but their number and their degrees $d_i := \deg \sigma_i$ are unique. Let us set

$$d := \max_{1\le i\le n} d_i$$

and consider the *orbit map* $\sigma = (\sigma_1, \dots, \sigma_n) : V \to \mathbb{R}^n$. The image $\sigma(V)$ is a semialgebraic set in the categorical quotient $V/\!\!/G := \{y \in \mathbb{R}^n : P(y) = 0 \text{ for all } P \in \mathscr{I}\}$, where $\mathscr{I}$ is the ideal of relations between $\sigma_1, \dots, \sigma_n$. Since G is compact, σ is proper and separates orbits of G, and it thus induces a homeomorphism between the orbit space V/G and $\sigma(V)$.

Note that, by the differentiable slice theorem, V/G is a local model for the orbit space of differentiable G-manifolds.

The given data, i.e., the representation $G \circlearrowleft V$, a fixed system of basic invariants $\{\sigma_i\}_{i=1}^n$ with degrees $d_i = \deg \sigma_i$, and the corresponding orbit map σ, will be abbreviated by the triple $(G \circlearrowleft V, d, \sigma)$.

In this setting, we may consider the following lifting problem: let $f : U \to \mathbb{R}^n$, $U \subseteq \mathbb{R}^m$, be a map with some regularity (C^ω, C^∞, C^p, etc.) and image contained in $\sigma(V)$; for brevity we will write $f \in \mathscr{C}(f, \sigma(V))$, where $\mathscr{C}$ is the respective regularity class. How regular can a lifting $\overline{f} : U \to V$ of f be?

$$\begin{tikzcd} & V \arrow[d, "\sigma"] \\ U \arrow[ur, dashed, "\overline{f}"] \arrow[r, "f"'] & \sigma(V) \end{tikzcd}$$

Note that any two systems of basic invariants differ by a polynomial diffeomorphism. Thus this question is independent of the choice of $\{\sigma_i\}_{i=1}^n$ and hence of σ.

5.2.1. *Differentiable lifting.* The results known for this lifting problem generalize the results for hyperbolic polynomials. The role of the polynomial degree (as an order of differentiablity of the coefficients needed to guarantee certain regularity properties of the roots) is here played by d, the maximal degree of a minimal system of basic invariants.

Theorem 5.1 ([56, Proposition 3.1 and Theorem 4.4])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a real finite dimensional orthogonal representation of a compact Lie group. Let $I \subseteq \mathbb{R}$ be an open interval. Then:*

(1) *Each continuous curve $c : I \to \sigma(V)$ has a continuous lifting $\overline{c} : I \to V$.*
(2) *Each C^d curve $c : I \to \sigma(V)$ has a differentiable lifting $\overline{c} : I \to V$.*

In Parusiński and Rainer [82], the proof of Bronshtein's theorem presented in Section 3.2 was generalized and led to the following essentially optimal results.

The next theorem is the analogue of Bronshtein's theorem 3.2.

Theorem 5.2 ([82, Theorem 1])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a real finite dimensional orthogonal representation of a compact Lie group. Let $I \subseteq \mathbb{R}$ be an open interval. Each curve $c \in C^{d-1,1}(I, \sigma(V))$ has a lifting $\overline{c} \in C^{0,1}(I, V)$ such that, for any pair of relatively compact open intervals $I_0 \Subset I_1 \Subset I$,*

$$\operatorname{Lip}_{I_0}(\overline{c}) \le C \max_{1\le i\le n} \|c_i\|^{1/d_i}_{C^{d-1,1}(\overline{I}_1)}. \tag{40}$$

The constant $C > 0$ depends only on I_0, I_1, and the isomorphism classes of the representation $G \circlearrowleft V$ and respective minimal systems of basic invariants.

In analogy to Theorem 3.9, we have

Theorem 5.3 ([82, Theorem 2])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a real finite dimensional orthogonal representation of a compact Lie group. Let $I \subseteq \mathbb{R}$ be an open interval. Each curve $c \in C^d(I, \sigma(V))$ has a lifting $\overline{c} \in C^1(I, V)$.*

For finite groups and polar representations, a precursor of this theorem was obtained by [57, 59] (by reduction to the case $\mathrm{S}_d \circlearrowleft \mathbb{R}^d$ and thus to Bronshtein's theorem).

An orthogonal representation $G \circlearrowleft V$ is called *polar*, if there exists a linear subspace $\Sigma \subseteq V$, called a *section* or a *Cartan subspace*, which meets each orbit orthogonally; cf. [28], [29]. The trace of the G-action on Σ is the action of the *generalized Weyl group* $W(\Sigma) = N_G(\Sigma)/Z_G(\Sigma)$ on Σ, where $N_G(\Sigma) := \{g \in G : g\Sigma = \Sigma\}$ and $Z_G(\Sigma) := \{g \in G : gs = s \text{ for all } s \in \Sigma\}$. This group is finite, and it is a reflection group if G is connected. The algebras $\mathbb{R}[V]^G$ and $\mathbb{R}[\Sigma]^{W(\Sigma)}$ are isomorphic via restriction, by a generalization of Chevalley's restriction theorem, due to [29] and independently [110], and thus the orbit spaces V/G and $\Sigma/W(\Sigma)$ are isomorphic. Consequently, the lifting problem can be reduced to a finite group action.

For finite group actions, Theorem 5.2 and Theorem 5.3 can be sharpened.

Theorem 5.4 ([82, Corollary 1 and 3])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a real finite dimensional orthogonal representation of a finite group. Let $I \subseteq \mathbb{R}$ be an open interval. Then:*

1. *Any continuous lifting $\overline{c}$ of $c \in C^{d-1,1}(I, \sigma(V))$ is locally Lipschitz and satisfies (40).*
2. *Any differentiable lifting $\overline{c}$ of $c \in C^d(I, \sigma(V))$ is C^1.*
3. *Each curve $c \in C^{2d}(I, \sigma(V))$ has a C^1 lifting whose second order derivatives exist everywhere.*

If G is a compact Lie group of positive dimension, then a Lipschitz (or a C^1) lifting $\overline{c}$ of c can be distorted to a continuous lifting that is not locally Lipschitz simply by taking $g\overline{c}$, where $g : I \to G$ is a suitable continuous curve.

As a consequence of the uniformity of Theorem 5.2, we obtain a lifting result for maps in several variables.

Theorem 5.5 ([82, Corollary 2])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a real finite dimensional orthogonal representation of a finite group. Let $U \subseteq \mathbb{R}^m$ be open and $f \in C^{d-1,1}(U, \sigma(V))$. If $\overline{f} \in C^0(\Omega, V)$ is a continuous lifting of f on an open subset $\Omega \subseteq U$, then $\overline{f}$ is locally Lipschitz and, for any pair of relatively compact open subsets $\Omega_0 \Subset \Omega_1 \Subset \Omega$,*

$$\operatorname{Lip}_{\Omega_0}(\overline{f}) \le C \max_{1 \le i \le n} \|f_i\|^{1/d_i}_{C^{d-1,1}(\overline{\Omega}_1)}. \tag{41}$$

The constant $C > 0$ depends only on Ω_0, Ω_1, m, and on the isomorphism classes of the representation $G \circlearrowleft V$ and respective minimal systems of basic invariants.

Not every representation admits unrestricted continuous lifting. For instance, the orbit space of a finite rotation group acting in the standard way on $\mathbb{R}^2$ is homeomorphic to the set C obtained from the sector $\{re^{i\varphi} \in \mathbb{C} : r \in [0, \infty), 0 \le$

$\varphi \le \varphi_0\}$ by identifying the rays that constitute its boundary. A loop on C cannot be lifted to a loop in $\mathbb{R}^2$ unless it is homotopically trivial in $C \setminus \{0\}$. Reflection groups admit continuous lifting defined everywhere, e.g., the hyperbolic case $\mathrm{S}_d \circlearrowleft \mathbb{R}^d$.

In view of the examples for hyperbolic polynomials, these lifting results are generally optimal.

5.2.2. *Real analytic lifting.* Real analytic curves can always be lifted.

Theorem 5.6 ([3, Lemma 3.8 and Theorem 4.2], [82, Theorem 4])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a real finite dimensional orthogonal representation of a compact Lie group. Let $I \subseteq \mathbb{R}$ be an open interval. Each curve $c \in C^\omega(I, \sigma(V))$ has a lifting $\bar{c} \in C^\omega(I, V)$.*

That there always exist local real analytic liftings was shown in [3]. As observed in [82], the local liftings can be glued to a global real analytic lifting, since the first Čech cohomology group $\check{H}^1(I, G^a)$ vanishes, where G^a denotes the sheaf of real analytic maps $I \supseteq U \to G$, by [111].

Similarly, one sees that *generic* C^∞ curves admit liftings. The genericity condition is a generalization of normal nonflatness considered for polynomials in Section 3.3.3: for integer $s \ge 0$ let A_s be the union of all orbit type strata S of V/G (cf. Section 5.3.4) of dimension $\dim S \le s$ and let $I_s \subseteq \mathbb{R}[V]^G$ be the ideal of invariant polynomials vanishing on A_{s-1}. Now $c \in C^\infty(I, \sigma(V))$ is called *normally nonflat* if, for each $t_0 \in I$, there exists $f \in I_s$ such that the Taylor series of $f \circ c$ at t_0 is not zero, where s is minimal with the property that the germ of c at t_0 is contained in A_s. Roughly speaking, it means that lower dimensional orbit type strata are not met with infinite order of flatness.

Theorem 5.7 ([3, Theorem 4.1])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a real finite dimensional orthogonal representation of a compact Lie group. Let $I \subseteq \mathbb{R}$ be an open interval. Each normally nonflat curve $c \in C^\infty(I, \sigma(V))$ has a lifting $\bar{c} \in C^\infty(I, V)$.*

In the C^∞ case, the gluing of local liftings is much easier; cf. [3, Lemma 3.8].

Real analytic maps depending on several variables cannot be lifted in general, but they admit liftings after composition with local blowings-up (i.e., blowings-up over an open subset composed with the inclusion of this subset).

Theorem 5.8 ([93, Theorem 5.4])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a real finite dimensional orthogonal representation of a compact Lie group. Let M be a real analytic manifold and $f \in C^\omega(M, \sigma(V))$. For any compact subset $K \subseteq M$ there exist a neighborhood U of K and a finite covering $\{\pi_j : U_j \to U\}$ of U, where each π_j is a composite of finitely many local blowings-up with smooth center, such that, for all j, the map $f \circ \pi_j$ has a real analytic lifting on U_j.*

Actually, this continues to hold in suitable quasianalytic classes of C^∞ functions (instead of C^ω); cf. [93].

5.3. **The main tools.** Let us briefly discuss the main tools used for the lifting problem and identify their counterpart in the study of hyperbolic polynomials.

5.3.1. *Removing fixed points.* Let $V^G := \{v \in V : Gv = v\}$ be the linear subspace of fixed points and let V' be its orthogonal complement in V. Then $V = V^G \oplus V'$, $\mathbb{R}[V]^G = \mathbb{R}[V^G] \otimes \mathbb{R}[V']^G$, and $V/G = V^G \times V'/G$. Thus it suffices to study the lifting problem for $G \circlearrowleft V'$ or equivalently to assume $V^G = \{0\}$. This corresponds

to the Tschirnhausen transformation (see Section 2.2.1); indeed, $(\mathbb{R}^d)^{\mathrm{S}_d} = \{x \in \mathbb{R}^d : x_1 = x_2 = \cdots = x_d\}$ and $(\mathbb{R}^d)' = \{x \in \mathbb{R}^d : x_1 + x_2 + \cdots + x_d = 0\}$.

5.3.2. *Dominant invariant.* We may assume that $v \mapsto \langle v \mid v\rangle = \|v\|^2$ belongs to the system of basic invariants $\{\sigma_i\}_{i=1}^n$, say it is σ_1. This does not change d; in fact, $V^G = \{0\}$ implies $d \geq 2$. Then σ_1 is dominant in the sense that

$$|\sigma_i|^{1/d_i} \leq C(\sigma)\,|\sigma_1|^{1/d_1}, \quad 1 \leq i \leq n,$$

by homogeneity. This corresponds to the dominance of the second coefficient of monic hyperbolic polynomials in Tschirnhausen form (see Lemma 2.4).

5.3.3. *The slice theorem.* For $v \in V$, let $N_v := T_v(Gv)^\perp$ be the normal subspace of the orbit Gv at v. It carries a natural action $G_v \circlearrowleft N_v$ of the *isotropy subgroup* $G_v := \{g \in G : gv = v\}$. The crossed product (or associated bundle) $G \times_{G_v} N_v$ carries the structure of an affine real algebraic variety as the categorical (and geometric) quotient $G \times N_v /\!\!/ G_v$ with respect to the action $G_v \circlearrowleft (G \times N_v)$, $h(g,x) := (gh^{-1}, hx)$. The G-equivariant polynomial mapping $\phi : G \times_{G_v} N_v \to V$, $[g,x] \mapsto g(v+x)$, induces a polynomial mapping $\psi : (G \times_{G_v} N_v)/\!\!/G \to V/\!\!/G$ sending $(G \times_{G_v} N_v)/G$ into V/G. The G_v-equivariant embedding $\alpha : N_v \hookrightarrow G \times_{G_v} N_v$, $x \mapsto [e,x]$, induces an isomorphism $\beta : N_v/\!\!/G_v \to (G \times_{G_v} N_v)/\!\!/G$ mapping N_v/G_v onto $(G \times_{G_v} N_v)/G$.

$$\begin{tikzcd}
N_v \arrow[r, hook, "\alpha"] \arrow[rr, bend left, "\eta"] \arrow[d, "\tau"'] & G \times_{G_v} N_v \arrow[r, "\phi"] \arrow[d] & V \arrow[d, "\sigma"] \\
N_v/G_v \arrow[r] \arrow[d, hook] & (G \times_{G_v} N_v)/G \arrow[r] \arrow[d, hook] & V/G \arrow[d, hook] \\
N_v/\!\!/G_v \arrow[r, "\beta"] \arrow[rr, bend right, "\theta"'] & (G \times_{G_v} N_v)/\!\!/G \arrow[r, "\psi"] & V/\!\!/G
\end{tikzcd}$$

Theorem 5.9 ([72], [102])**.** *There is an open ball $B_v \subseteq N_v$ centered at the origin such that the restriction of ϕ to $G \times_{G_v} B_v$ is an analytic G-isomorphism onto a G-invariant neighborhood of v in V. The mapping θ is a local analytic isomorphism at 0 which induces a local homeomorphism of N_v/G_v and V/G.*

The slice theorem allows to reduce the lifting problem locally to the slice representation $G_v \circlearrowleft N_v$. This reduction replaces the splitting principle (see Lemma 2.5 and Section 3.2.1) for polynomials. (The passage to the slice representation at $v = (\lambda_1, \ldots, \lambda_d)$ for the standard action $\mathrm{S}_d \circlearrowleft \mathbb{R}^d$ corresponds to a *full* splitting $P_{\tilde{a}} = P_{b^1} P_{b^2} \cdots P_{b^k}$, where each factor P_{b^i} corresponds to precisely one of the distinct elements among the $\lambda_1, \ldots, \lambda_d$ and $\deg P_{b^i}$ is its multiplicity.)

The reduction and the assumption $V^G = \{0\}$ allow to proceed by induction on the "size" of G: if H and G are compact Lie groups, then H is *smaller* than G if either $\dim H < \dim G$ or, given that $\dim H = \dim G$, H has fewer connected components than G. For finite groups this reduces to induction on the group order.

5.3.4. *Orbit type stratification.* The conjugacy class (H) of $H = G_v$ in G is called the *type* of the orbit Gv. Let $V_{(H)}$ be the union of all orbits of type (H). The collection of the connected components of the smooth manifolds $V_{(H)}/G$ forms the *orbit type stratification of* V/G; cf. [102]. By [13], there is an identification between the orbit type stratification of V/G and the natural stratification of $\sigma(V)$ as a semialgebraic set (for the construction of the latter see [13, p. 246] or [69, Section 25]); it is analytically locally trivial and thus satisfies Whitney's conditions (a) and (b). The inclusion relation on the set of subgroups of G induces a partial ordering on the family of orbit types. There is a unique minimal orbit type, the *principal orbit type*, corresponding to the open and dense submanifold V_{reg} consisting of points v, where the slice representation $G_v \circlearrowleft N_v$ is trivial. The projection $V_{\mathrm{reg}} \to V_{\mathrm{reg}}/G$ is a locally trivial fiber bundle. There are only finitely many isomorphism classes of slice representations.

5.4. **Examples and applications.**

5.4.1. *Differentiable eigenvalues of real symmetric matrices.* Let the orthogonal group $\mathrm{O}(d) = \mathrm{O}(\mathbb{R}^d)$ act by conjugation on the real vector space $\mathrm{Sym}(d)$ of real symmetric $d \times d$ matrices,

$$\mathrm{O}(d) \times \mathrm{Sym}(d) \ni (S, A) \mapsto SAS^{-1} = SAS^t \in \mathrm{Sym}(d).$$

The algebra of invariant polynomials $\mathbb{R}[\mathrm{Sym}(d)]^{\mathrm{O}(d)}$ is isomorphic to $\mathbb{R}[\mathrm{Diag}(d)]^{\mathrm{S}_d}$ by restriction, where $\mathrm{Diag}(d)$ is the vector space of real diagonal $d \times d$ matrices upon which S_d acts by permuting the diagonal entries. More precisely, $\mathbb{R}[\mathrm{Sym}(d)]^{\mathrm{O}(d)} = \mathbb{R}[\Sigma_1, \dots, \Sigma_d]$, where

$$\Sigma_i(A) = \mathrm{Trace}(\textstyle\bigwedge^i A : \bigwedge^i \mathbb{R}^d \to \bigwedge^i \mathbb{R}^d)$$

is the i-th characteristic coefficient of A and $\Sigma_i|_{\mathrm{Diag}(d)} = \sigma_i$, where σ_i is the i-th elementary symmetric polynomial and we identify $\mathrm{Diag}(d) \cong \mathbb{R}^d$ (cf. [76, 7.1]). This means that the representation $\mathrm{O}(d) \circlearrowleft \mathrm{Sym}(d)$ is polar and $\mathrm{Diag}(d)$ forms a section.

Let $A : \mathbb{R} \to \mathrm{Sym}(d)$ be a curve of symmetric matrices of some regularity. The characteristic polynomial χ_A is a curve of monic hyperbolic polynomials of the same regularity. In this case, as we already know from Section 3.5, the regularity results provided by the general theory for hyperbolic polynomials are not optimal.

A heuristic explanation is that in this problem only a "partial" lifting is necessary: the curve χ_A in the orbit space is the projection of the curve A under $\mathrm{Sym}(d) \to \mathrm{Sym}(d)/\mathrm{O}(d)$ and is then lifted over $\mathrm{Diag}(d) \to \mathrm{Diag}(d)/\mathrm{S}_d$.

$$\begin{tikzcd}
 & & \mathbb{R} \arrow[dll, dashed, bend right] \arrow[dl, "A"'] \arrow[dd, "\chi_A"] & \\
\mathrm{Diag}(d) \arrow[r, hook] \arrow[d] & \mathrm{Sym}(d) \arrow[d] & & \\
\mathrm{Diag}(d)/\mathrm{S}_d \arrow[r, equal] & \mathrm{Sym}(d)/\mathrm{O}(d) \arrow[r, equal] & \sigma(\mathbb{R}^d) \arrow[r, hook] & \mathbb{R}^d
\end{tikzcd}$$

5.4.2. *Differentiable decomposition of nonnegative functions into sums of squares.* Let the orthogonal group $\mathrm{O}(n)$ act in the standard way on $\mathbb{R}^n$. The algebra of invariant polynomials $\mathbb{R}[\mathbb{R}^n]^{\mathrm{O}(n)}$ is generated by $\sigma = \sum_{i=1}^n x_i^2$; note that $d = 2$.

The orbit space $\mathbb{R}^n/\operatorname{O}(n)$ can be identified with the half-line $[0,\infty)=\sigma(\mathbb{R}^n)$. Each line through the origin of $\mathbb{R}^n$ forms a section of $\operatorname{O}(n)\circlearrowleft\mathbb{R}^n$.

Lifting a nonnegative function f over σ means decomposing f into a sum of n squares. We conclude (from Theorem 5.5) that any nonnegative $C^{1,1}$ function $f:\mathbb{R}^m\to[0,\infty)$ is the square of a $C^{0,1}$ function g. The image of the lifting $(g,0,\dots,0)$ lies in the section $\mathbb{R}(1,0,\dots,0)$ of $\operatorname{O}(n)\circlearrowleft\mathbb{R}^n$.

There are stronger results that benefit from the additionally available space:

Theorem 5.10 ([34])**.** *Any nonnegative $C^{3,1}$ function $f:\mathbb{R}^m\to[0,\infty)$ is a sum of $n=n(m)$ squares of $C^{1,1}$ functions.*

This is sharp in the sense that there exist C^∞ functions $f:\mathbb{R}^m\to[0,\infty)$, for $m\geq 4$, that are not sums of squares of C^2 functions; see [18]. In fact, a real nonnegative homogeneous polynomial p of degree $2d$ on $\mathbb{R}^m$ that is a sum of squares of C^d functions f_i is necessarily a sum of squares of polynomials, since Taylor expansion of f_i gives $f_i(x)=q_i(x)+o(|x|^d)$ at $x=0$, where q_i is a homogeneous polynomial of degree d, so that $p=\sum_i q_i^2$. But there exist nonnegative homogeneous polynomials of degree 4 on $\mathbb{R}^4$ that are not sums of squares of polynomials; cf. [16, Section 6.3]. For the same reason, there are nonnegative C^∞ functions on $\mathbb{R}^3$ that are not sums of squares of C^3 functions.

In dimension $m=1$ (where there are no algebraic obstructions) there is the following result.

Theorem 5.11 ([17])**.** *Let $p\in\mathbb{N}$. Any nonnegative C^{2p} function $f:\mathbb{R}\to[0,\infty)$ is the sum of two squares of C^p functions.*

The decomposition depends on p.

5.4.3. *Polynomials with symmetries.* In [71] the lifting problem was used to improve upon the regularity problem for hyperbolic polynomials with symmetries. The idea is the following. A curve of monic hyperbolic polynomials of degree d is a curve p in the semialgebraic set $\operatorname{Hyp}(d)=\sigma(\mathbb{R}^d)\subseteq\mathbb{R}^d$, where σ consists of the elementary symmetric functions σ_i. If p lies in some proper semialgebraic subset of $\operatorname{Hyp}(d)$, then p has to fulfill more constraints and thus, in general, the conditions that guarantee the existence of differentiable systems of the roots are weaker. For instance, if we know that the roots of p fulfill some linear relations, i.e., they lie in some linear subspace V of $\mathbb{R}^d$, then p lies in $\sigma(V)\subseteq\operatorname{Hyp}(d)$. The symmetries of the roots are represented by the induced action $G\circlearrowleft V$ of

$$G:=\frac{\{\tau\in\mathrm{S}_d:\tau V=V\}}{\{\tau\in\mathrm{S}_d:\tau v=v\text{ for all }v\in V\}}.$$

The solution of the lifting problem can be brought to bear, if the restrictions $\sigma_i|_V$, $1\leq i\leq d$, generate $\mathbb{R}[V]^G$. Indeed, the orbit type stratification of $G\circlearrowleft V$ is generally coarser than the restriction to V of the orbit type stratification of $\mathrm{S}_d\circlearrowleft\mathbb{R}^d$.

5.5. **A reformulation of the regularity problem for general polynomials.** If we let the symmetric group S_d act on $\mathbb{C}^d$, by permuting the coordinates, then again $\mathbb{C}[\mathbb{R}^d]^{\mathrm{S}_d}$ is generated by the elementary symmetric functions σ_i, $1\leq i\leq d$. The map $\sigma=(\sigma_1,\dots,\sigma_d):\mathbb{C}^d\to\mathbb{C}^d$ is onto. Each point in $\mathbb{C}^d$ represents a monic polynomial of degree d with complex coefficients and the σ-fiber over this point is the S_d-orbit through the d-tuple of its roots (with multiplicities).

A family of monic polynomials $P_a(x)$, $x \in U \subseteq \mathbb{R}^m$, is a map $p : U \to \mathbb{C}^d$ and a system of roots for the family P_a is a lifting $\overline{p} : U \to \mathbb{C}^d$ of p over σ, i.e., $\sigma \circ \overline{p} = p$. Given that p is regular of some kind, how regular can a lifting of p be?

$$\begin{tikzcd} & \mathbb{C}^d \arrow[d, "\sigma"] \\ U \arrow[ur, dashed, "\overline{p}"] \arrow[r, "p"'] & \mathbb{C}^d \end{tikzcd}$$

Note that, in contrast to the hyperbolic case (cf. Section 5.1), there are no constraints for the map p, since $\sigma(\mathbb{C}^d) = \mathbb{C}^d$.

Let us see now how this problem can be studied in greater generality. From now on, representation spaces V will be complex and finite dimensional.

5.6. **Representations of linearly reductive groups.** An algebraic group G is called *linearly reductive* if for each rational representation V and each subrepresentation $W \subseteq V$ there is a subrepresentation $W' \subseteq V$ such that $V = W \oplus W'$.

By Hilbert's finiteness theorem, for rational representations V of linearly reductive groups G, the algebra of G-invariant polynomials $\mathbb{C}[V]^G$ is finitely generated. Let $\{\sigma_i\}_{i=1}^n$ be a system of homogeneous generators which we call *basic invariants*. In a *minimal* system of basic invariants the number n of elements σ_i and their degrees $d_i := \deg \sigma_i$ are uniquely determined; set

$$d := \max_{1 \le i \le n} d_i.$$

The map $\sigma = (\sigma_1, \dots, \sigma_n) : V \to \sigma(V) \subseteq \mathbb{C}^n$ can be identified with the morphism $V \to V /\!\!/ G$ induced by the inclusion $\mathbb{C}[V]^G \to \mathbb{C}[V]$; the *categorical quotient* $V /\!\!/ G$ is the affine variety with coordinate ring $\mathbb{C}[V]^G$. In this setting, $V /\!\!/ G$ is generally not a *geometric quotient*: the G-orbits in V are not in a one-to-one correspondence with the points in $V /\!\!/ G$. In fact, for every point $z \in V /\!\!/ G$ there is a unique closed orbit in the fiber $\sigma^{-1}(z)$ which lies in the closure of every other orbit in this fiber. On the other hand, if G is a finite group, then all G-orbits are closed and thus $V /\!\!/ G$ is a geometric quotient isomorphic to the orbit space V/G.

We write again $(G \circlearrowleft V, d, \sigma)$ for this setup. If $f : U \to \mathbb{C}^n$, $U \subseteq \mathbb{R}^m$, is a map with some regularity and image contained in $\sigma(V)$ (which in general is a proper subset of $\mathbb{C}^n$), we ask how regular a lifting $\overline{f}$ of f can be. By a *lifting* of f we mean a map $\overline{f} : U \to V$ such that $f = \sigma \circ \overline{f}$ and, for all $x \in U$, $\overline{f}(x)$ lies in the unique closed orbit in the fiber $\sigma^{-1}(f(x))$.

$$\begin{tikzcd} & V \arrow[d, "\sigma"] \\ U \arrow[ur, dashed, "\overline{f}"] \arrow[r, "f"'] & \sigma(V) \end{tikzcd}$$

The problem is independent of the choice of the basic invariants, since any two choices differ by a polynomial diffeomorphism.

The main difference to the real problem introduced in Section 5.2 is that there is no invariant polynomial that dominates all the others. Even for finite G, where we can always choose an invariant Hermitian inner product on V (by averaging over G) and hence assume that the representation is unitary, the invariant form $v \mapsto \|v\|^2$

is not a member of $\mathbb{C}[V]^G$. This makes the complex case much more difficult and causes a larger loss of regularity.

5.6.1. *Sobolev lifting.* Let us first assume that G is a finite group. Then each continuous curve $c : I \to \sigma(V)$, where $I \subseteq \mathbb{R}$ is an interval, has a continuous lifting $\overline{c} : I \to V$, by [70, Theorem 5.1]. In Parusiński and Rainer [86], the optimal Sobolev regularity for the roots of polynomials was extended to group representations. The following result generalizes Theorem 4.2.

Theorem 5.12 ([86, Theorem 1.1])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a complex finite dimensional representation of a finite group G. Let $c \in C^{d-1,1}([\alpha,\beta], \sigma(V))$ be a curve defined on an open bounded interval (α,β). Then each continuous lifting $\overline{c} : (\alpha,\beta) \to V$ of c over σ is absolutely continuous and belongs to $W^{1,p}((\alpha,\beta), V)$ with*

$$\|\overline{c}'\|_{L^p((\alpha,\beta))} \le C \max_{1\le j\le n} \|c_j\|^{1/d_j}_{C^{d-1,1}([\alpha,\beta])} \tag{42}$$

for all $1 \le p < d/(d-1)$, where C is a constant which depends only on the representation $G \circlearrowleft V$, the length of the interval (α,β), and p.

The bound (42), similarly as (32), is not scale invariant.

Remark 5.13. Let us comment on how the constant C in (42) depends on the length of the interval (α,β).

(a) In general, the constant C is of the form

$$C(G \circlearrowleft V, p) \max\{1, (\beta-\alpha)^{1/p}, (\beta-\alpha)^{-1+1/p}\}.$$

(b) If the curve c starts, ends, or passes through 0 (the most singular point in $\sigma(V)$), then C is of the form

$$C(G \circlearrowleft V, p) \max\{1, (\beta-\alpha)^{1/p}\}. \tag{43}$$

(c) If the representation is coregular, then again the constant is of the form (43). A representation $G \circlearrowleft V$ is called *coregular* if $\mathbb{C}[V]^G$ is isomorphic to a polynomial algebra, i.e., there is a system of basic invariants without polynomial relations among them. By the Shephard–Todd–Chevalley theorem ([104], [23], [103]), this is the case if and only if G is generated by pseudoreflections (i.e., invertible linear transformations of finite order and fixed point space a hyperplane).

(d) The constant is also of the form (43) if the curve c satisfies $c^{(j)}(\alpha) = c^{(j)}(\beta) = 0$ for all $j = 1, \dots, d-1$.

Note that, in view of the counterexamples for the regularity of the roots of polynomials, the results of Theorem 5.12 (as most of the subsequent results) are optimal.

The lifting of mappings defined in open domains of dimension $m > 1$ essentially admits the same regularity as for curves, provided that continuous lifting is possible. But, in general, there are topological obstructions for continuous lifting.

Theorem 5.14 ([86, Theorem 1.4])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a complex finite dimensional representation of a finite group G. Let $f \in C^{d-1,1}(\overline{\Omega}, \sigma(V))$, where $\Omega \subseteq \mathbb{R}^m$ is an open bounded box $\Omega = I_1 \times \cdots \times I_m$. Then each continuous lifting $\overline{f} : U \to V$ of f over σ defined on an open subset $U \subseteq \Omega$ belongs to $W^{1,p}(U, V)$ and satisfies*

$$\|\nabla \overline{f}\|_{L^p(U)} \le C(G \circlearrowleft V, \Omega, m, p) \max_{1\le j\le n} \|f_j\|^{1/d_j}_{C^{d-1,1}(\overline{\Omega})} \tag{44}$$

for all $1 \le p < d/(d-1)$.

The case $U = \Omega$ is not excluded. It is clear that Theorem 5.14 implies a version of the statement, where $\Omega \subseteq \mathbb{R}^m$ is any bounded open set, $U \Subset \Omega$ is relatively compact open in Ω, and the constant also depends on U (or more precisely on a cover of U by boxes contained in Ω). Concerning a global result we have the following.

Remark 5.15. If $G \circlearrowleft V$ is coregular, then Theorem 5.14 holds as stated for any bounded Lipschitz domain Ω; cf. [86, Remark 1.5].

If there is no continuous lifting, it is natural to ask how bad the discontinuities can be:

Open Problem 5.16. *When continuous lifting is impossible, do there exist liftings of bounded variation?*

See Theorem 5.23 for a partial answer.

In the general case, for an infinite linearly reductive group G, it is not clear if a continuous curve in $\sigma(V)$ admits a continuous lifting to V. The notion of *stability* in geometric invariant theory provides a remedy. A point $v \in V$ is called *stable* if the orbit Gv is closed and the isotropy group G_v is finite. The subset $V^s \subseteq V$ of stable points is G-invariant and open in V, and its image $\sigma(V^s)$ is open in $V /\!\!/ G \cong \sigma(V)$ (cf. [78, Proposition 5.15]). The restriction $\sigma : V^s \to \sigma(V^s)$ of the map σ provides a one-to-one correspondence between points in $\sigma(V^s) \cong V^s/G$ and G-orbits in V^s, that is V^s/G is a geometric quotient. A continuous curve $c : I \to \sigma(V^s)$ has a continuous lifting $\overline{c} : I \to V^s$, cf. [86, Lemma 1.6]. Theorem 5.12 has the following corollary.

Corollary 5.17 ([86, Theorem 1.7])**.** *Let* $(G \circlearrowleft V, d, \sigma)$ *be a rational complex finite dimensional representation of a linearly reductive group* G. *Let* $c \in C^{d-1,1}([\alpha,\beta], \sigma(V^s))$ *be a curve defined on a compact interval* $[\alpha,\beta]$ *with* $c([\alpha,\beta]) \subseteq \sigma(V^s)$. *Then there exists an absolutely continuous lifting* $\overline{c} : [\alpha,\beta] \to V^s$ *of* c *over* σ *which belongs to* $W^{1,p}([\alpha,\beta], V^s)$ *with*

$$\|\overline{c}'\|_{L^p([\alpha,\beta])} \le C(G \circlearrowleft V, [\alpha,\beta], p) \max_{1\le j\le n} \|c_j\|^{1/d_j}_{C^{d-1,1}([\alpha,\beta])} \tag{45}$$

for all $1 \le p < d/(d-1)$.

For a mapping f defined on a compact subset K of $\mathbb{R}^m$ with $f(K) \subseteq \sigma(V^s)$ the situation is more complicated: we can apply Theorem 5.14 to the slice representations at any point $v \in V^s$, but it is not clear if these local (and partial) lifts can be glued in a continuous fashion.

More can be said for polar representations (which include e.g. the adjoint actions). We encountered polar representations already in Section 5.2 in the real setup. Let us briefly discuss them in the framework of rational representations V of linearly reductive groups G; we follow [29]. Let $v \in V$ be such that the orbit Gv is closed and consider the linear subspace $\Sigma_v = \{x \in V : \mathfrak{g}x \subseteq \mathfrak{g}v\}$, where $\mathfrak{g}$ is the Lie algebra of G. All orbits that intersect Σ_v are closed, whence $\dim \Sigma_v \le \dim V /\!\!/ G$. The representation $G \circlearrowleft V$ is said to be *polar* if there exists $v \in V$ with closed orbit Gv and $\dim \Sigma_v = \dim V /\!\!/ G$. Then Σ_v is called a *section* or *Cartan subspace* of V. Any two sections are G-conjugate. Let us fix a section Σ. All closed orbits in V intersect Σ. The *Weyl group* $W := N_G(\Sigma)/Z_G(\Sigma)$, where $N_G(\Sigma) = \{g \in G : g\Sigma = \Sigma\}$

is the normalizer and $Z_G(\Sigma) = \{g \in G : gx = x \text{ for all } x \in \Sigma\}$ is the centralizer of Σ in G, is finite and the intersection of any closed G-orbit in V with Σ is precisely one W-orbit. The ring $\mathbb{C}[V]^G$ is isomorphic via restriction to the ring $\mathbb{C}[\Sigma]^W$. If G is connected, then W is a pseudoreflection group and hence $\mathbb{C}[V]^G \cong \mathbb{C}[\Sigma]^W$ is a polynomial ring, by the Shephard–Todd–Chevalley theorem.

Corollary 5.18 ([86, Theorem 1.8])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a polar representation of a linearly reductive group G. Then:*

(1) *Each curve $c \in C^{d-1,1}([\alpha,\beta], \sigma(V))$ has an absolutely continuous lifting $\overline{c} : (\alpha,\beta) \to V$ of c over σ which belongs to $W^{1,p}((\alpha,\beta), V)$ for all $1 \le p < d/(d-1)$ and satisfies (42).*
(2) *Let $f \in C^{d-1,1}(\overline{\Omega}, \sigma(V))$, where $\Omega \subseteq \mathbb{R}^m$ is an open bounded box $\Omega = I_1 \times \cdots \times I_m$. Each continuous lifting $\overline{f}$ defined on an open subset $U \subseteq \Omega$ with values in a section Σ is of class $W^{1,p}$ on U for all $1 \le p < d/(d-1)$ and satisfies (44).*
(3) *If G is connected, then the constant in (42) is of the form (43) and Ω can be any bounded Lipschitz domain.*

This follows from applying Theorem 5.12 and Theorem 5.14 to $W \circlearrowleft \Sigma$.

Open Problem 5.19. *To what extent are the above results true for general (non-polar) representation?*

5.6.2. *Analytic lifting.* Let $(G \circlearrowleft V, d, \sigma)$ be a complex finite dimensional representation of a finite group G. Then the orbit space coincides with the categorical quotient $V /\!\!/ G$ which is a normal affine variety. Thus the orbit space has the natural structure of a complex analytic set and there are several types of morphisms into $V /\!\!/ G$, like regular, rational, or holomorphic. In [58] those regular, holomorphic, or formal maps into $V /\!\!/ G$ that admit a regular, holomorphic, or formal lifting to V are characterized. The conditions are formulated in terms of the orbit type stratification of the orbit space which, in this case (compare with Section 5.3.4), is finer than its stratification as affine variety and with the use of spaces of jets at $0 \in \mathbb{C}^m$ of morphisms $\mathbb{C}^m \to V /\!\!/ G$ of finite and infinite order.

Now let $(G \circlearrowleft V, d, \sigma)$ be a rational representation of a linearly reductive group G. Suppose that $c : \mathbb{R} \to V /\!\!/ G \cong \sigma(V) \subseteq \mathbb{C}^n$ is real analytic. In contrast to the real case, there does not always exist a real analytic lifting of c over σ. But there is the following generalization of Puiseux's theorem.

Theorem 5.20 ([70, Theorem 3.3 and 3.4])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a rational representation of a linearly reductive group G. Then:*

(1) *Let $I \subseteq \mathbb{R}$ be an open interval, $t_0 \in I$, and $c \in C^\omega(I, \sigma(V))$. Then there exists a positive integer N such that, locally near t_0, $t \mapsto c(t_0 \pm (t-t_0)^N)$ admits a real analytic lifting $\overline{c}_\pm$ to V.*
(2) *Let $U \subseteq \mathbb{C}$ be open and connected, $z_0 \in U$, and $c \in \mathcal{H}(U, \sigma(V))$, i.e., c is holomorphic. Then there exists a positive integer N such that, locally near z_0, $z \mapsto c(z_0 + (z-z_0)^N)$ admits a holomorphic lifting $\overline{c}$ to V.*

In analogy to Section 5.2.2, one defines that c is normally nonflat at t_0, see [70, Section 2.5]; roughly speaking, it means that c does not meet lower dimensional strata of the orbit type stratification of $V /\!\!/ G$ with infinite order of flatness at t_0. Then in (1) we may take $c \in C^\infty(I, \sigma(V))$ if c is normally nonflat at t_0 and conclude

that $\overline{c}_\pm$ is of class C^∞. Or we may work with suitable quasianalytic subclasses of C^∞.

As a consequence, we may conclude that real analytic curves in $\sigma(V)$ have liftings that are locally absolutely continuous.

Theorem 5.21 ([70, Theorem 5.4])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a rational representation of a linearly reductive group G. Let $I \subseteq \mathbb{R}$ be an open interval. Each $c \in C^\omega(I, \sigma(V))$ admits a lifting $\overline{c} : I \to V$ that is locally absolutely continuous.*

Again this remains true for smooth curves in $\sigma(V)$ that are normally nonflat.

For maps in several variables into $\sigma(V)$ there is an analogue of Theorem 5.8, where, however, local blowings-up *and* local power substitutions must be used. A *local power substitution* is the composite of the inclusion of a coordinate chart U and a map $W \to U$ given in local coordinates by

$$(x_1, \dots, x_m) \mapsto ((-1)^{\epsilon_1} x_1^{\gamma_1}, \dots, (-1)^{\epsilon_m} x_m^{\gamma_m}),$$

for some integers $\gamma_i \geq 1$ and $\epsilon_i \in \{0, 1\}$.

Theorem 5.22 ([93, Theorem 4.6])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a rational representation of a linearly reductive group G. Let M be a real analytic manifold and $f \in C^\omega(M, \sigma(V))$. For each compact $K \subseteq M$, there exist a neighborhood U of K in M and a finite covering $\{\pi_j : U_j \to U\}$ of U, where each π_j is a composite of finitely many maps each of which is either a local blowing-up with smooth center or a local power substitution, such that, for all j, the map $f \circ \pi_j$ has a real analytic lifting on U_j.*

The theorem remains true for suitable quasianalytic subclasses of C^∞ as well as for holomorphic maps. In the latter case, a local power substitution is simply a map of the form $(z_1, \dots, z_m) \mapsto (z_1^{\gamma_1}, \dots, z_m^{\gamma_m})$.

It is possible to extract information about the existence of weak liftings of f (without modification of its domain). For instance, there is the following result which gives a partial answer to Open Problem 5.16.

Theorem 5.23 ([93, Theorem 6.11])**.** *Let $(G \circlearrowleft V, d, \sigma)$ be a rational representation of a linearly reductive group G. Let $U \subseteq \mathbb{R}^m$ be open and $f \in C^\omega(U, \sigma(V))$. For each compact $K \subseteq U$, there exist a neighborhood W of K in U and a lifting $\overline{f} : W \to V$ of f of class SBV (i.e., special functions of bounded variation).*

Actually, f can be taken in a suitable quasianalytic class.

5.7. **Some remarks on the proofs.** We focus on the proof of Theorem 5.12. The foundation of the proof is that the result holds for finite rotation groups $C_d \cong \mathbb{Z}/d\mathbb{Z}$ acting in the standard way on $\mathbb{C}$, where $\mathbb{C}[\mathbb{C}]^{C_d}$ is generated by $z \mapsto z^d$ and a lifting of a map f is a solution of the radical equation $z^d = f$. This follows from Theorem 4.1. Among all representations of finite groups G of a fixed order $|G|$, it is the one with the worst loss of regularity, since in general $d \leq |G|$, by Noether's degree bound, and equality can only happen for cyclic groups.

The strategy in the general case is similar to the one described in Section 5.3. Evidently, one may reduce to the case that the linear subspace V^G of invariant vectors is trivial. Then Luna's slice theorem (see [72] or [102, Theorem 5.3]) allows us to reduce the problem locally to the slice representation $G_v \circlearrowleft N_v$ of the isotropy group G_v on N_v, where $T_v V \cong T_v(Gv) \oplus N_v$ is a G_v-splitting. Since in our case G is

finite, we have $N_v \cong V$. The assumption $V^G = \{0\}$ entails that for all $v \in V \setminus \{0\}$ the isotropy group G_v is a proper subgroup of G which suggests to use induction on the group order. For this induction scheme to work it is important that the slice reduction is uniform in the sense that it does not depend on the parameter t of the curve c in $\sigma(V) \subseteq \mathbb{C}^n$.

The lack of a dominant invariant polynomial (in contrast to the real case) forces one to choose at points t_0 with $c(t_0) \neq 0$ a *dominant* component c_k of c such that

$$|c_k^{1/d_k}(t_0)| = \max_{1\le j\le n} |c_j^{1/d_j}(t_0)|$$

and to work on small neighborhoods of t_0 with this distinguished component c_k.

The better bound in the case that $G \circlearrowleft V$ is coregular (cf. Remark 5.13; see also Remark 5.15 and Corollary 5.18(3)) is achieved by an application of Whitney's extension theorem to extend the curve c to a larger interval on which it can be suitably modified without changing it on the original interval (α, β). Coregularity guarantees that the extended curve still lies in $\sigma(V)$ (because $\sigma(V) = \mathbb{C}^n$).

6. Applications

6.1. **Zero sets of smooth functions.** It is well-known that every closed subset of $\mathbb{R}^n$ is the zero set Z_f of some real valued C^∞ function $f : \mathbb{R}^n \to \mathbb{R}$. So in general Z_f can be arbitrarily irregular. On the other hand, constraints for the derivatives of f entail regularity properties of Z_f; think of the implicit function theorem or of the solutions of elliptic and parabolic PDEs.

Let us assume here that our function f vanishes to some finite order γ at a point x_0. Then Malgrange's preparation theorem puts one in the position to study Z_f in a neighborhood of x_0, by using the regularity theorems for the roots of polynomials, that is Theorem 4.2 or Theorem 4.9. Indeed, in suitable local coordinates at x_0, we have $f = p \cdot u$, where p is a monic polynomial in one of the coordinates with C^∞ coefficients in the other coordinates and u is a C^∞ unit.

This setting is widely applicable, for instance, to solutions of second order elliptic equations, Laplace eigenfunctions and finite linear combinations of such ([32], [68], [50]), etc. Often Z_f is called the *nodal set* of f, but we will stick to the term *zero set*.

The following result is taken from Bär [8].

Theorem 6.1 ([8, Lemma 3])**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^∞ function that vanishes to order γ at x_0. Then there exist $r_0 > 0$ and a affine hyperplane H in $\mathbb{R}^n$ such that $Z_f \cap B(x_0, r_0)$ is contained in the union of countably many graphs of C^∞ functions from $H \cap B(x_0, r_0)$ to $H^\perp$. Moreover, for all $r < r_0$,*

$$\mathcal{H}^{n-1}(Z_f \cap B(x_0, r)) \le n 2^{n-1} \gamma\, r^{n-1}.$$

Note that Yomdin [116] obtained a similar result; he assumed that the partial derivatives of some order j of f are "small" in a certain precise sense and concluded that f behaves in many regards like a polynomial of degree $j-1$, in particular, it has vanishing order at most $j-1$.

The following theorem is a consequence of Theorem 4.2 or Theorem 4.9 (together with Malgrange's preparation theorem), due to Beck, Becker-Kahn, and Hanin [10].

Theorem 6.2 ([10, Theorem 1])**.** *Let $U \subseteq \mathbb{R}^n$ be open and let Θ be a compact smooth manifold (possibly with boundary). Let $f \in C^\infty(U \times \Theta)$ and write $f^\theta :=$*

$f(\cdot,\theta)$ for $\theta \in \Theta$. Suppose that f^{θ_0} has finite vanishing order γ at $x_0 \in U$. Then there exist $p > 1$, $r_0 > 0$, a neighborhood V of θ_0 in Θ, an affine hyperplane $H \subseteq \mathbb{R}^n$, and functions $f_i^\theta \in W^{1,p}(H, H^\perp)$, $1 \le i \le \gamma$, $\theta \in V$, with

$$\sup_{1\le i\le\gamma,\,\theta\in V} \|f_i^\theta\|_{W^{1,p}} < \infty$$

such that

$$Z_{f^\theta} \cap B(x_0, r_0) \subseteq \bigcup_{i=1}^{\gamma} \Gamma_{f_i^\theta}, \quad \theta \in V,$$

where $\Gamma_{f_i^\theta}$ is the graph of f_i^θ.

Remark 6.3. Theorem 6.2 was proved in [10] based on Theorem 4.9. Using the optimal result, Theorem 4.2, in the proof, one can get more precise information on p: the statement is true for all $1 \le p < \gamma/(\gamma-1)$.

By the same strategy, one gets similar parameterization results in the setting studied by Yomdin [116] (where the partial derivatives of some order are small) as well as for C^∞ functions all of whose derivatives are suitably controlled by a weight sequence (e.g. functions in a quasianalytic Denjoy–Carleman class), see Rainer [97]. In the latter case, the vanishing order is bounded (globally on a convex body) by a quantity computed from the weight sequence which has similarities with the polynomial degree.

Note that, if one is interested only in the properties of the zero set of a function near a point of finite vanishing order, then there is some freedom in the choice of the local coordinates (i.e., the hyperplane H in Theorem 6.2).

Open Problem 6.4. *Is is possible to improve the above results by choosing the local system of coordinates wisely?*

As a consequence of Theorem 6.1 or Theorem 6.2, one gets a nonconcentration estimate [10, Proposition 4]: suppose that the vanishing order of f^θ is at most γ for all $\theta \in \Theta$. For compact $K \subseteq U$ and m-rectifiable $E \subseteq \mathbb{R}^n$, for some $m \le n-1$, there exists $r_0 > 0$ such that

$$\mathcal{H}^{n-1}(Z_{f^\theta} \cap E_r' \cap K) \le C(n)\, r^{-1} \mathcal{H}^n(E_r), \quad r \le r_0,$$

for all $\theta \in \Theta$ and all images E' of Lipschitz maps $E \to K$ with Lipschitz constant ≤ 1, where E_r, E_r' denote the r-neighborhoods of E, E' in $\mathbb{R}^n$. This entails the following continuity result. Let $\mathrm{Sing}_f := \{x \in Z_f : \nabla f(x) = 0\}$.

Corollary 6.5 ([10, Corollary 2])**.** *Let $U \subseteq \mathbb{R}^n$ be open and $f \in C^\infty(U \times [0,1])$ such that $f^1 := f(\cdot, 1)$ has finite vanishing order on U. Then for any compact $K \subseteq U$ with $\mathcal{H}^{n-1}(K \cap \mathrm{Sing}_f) = 0$ we have*

$$\lim_{\theta\to 1} \mathcal{H}^{n-1}(Z_{f^\theta} \cap K) = \mathcal{H}^{n-1}(Z_{f^1} \cap K).$$

In particular, the authors of [10] conclude a continuity result for the zero sets of the heat flow on a compact smooth Riemannian manifold (M, g): if $u : M \times (0,\infty) \to \mathbb{R}$ solves

$$(\partial_t + \Delta_g) u(x,t) = 0$$

with initial data $u^0 \in L^2(M)$, then

$$\lim_{t\to 1} \mathcal{H}^{n-1}(Z_{u^t}) = \mathcal{H}^{n-1}(Z_\psi),$$

where ψ is the first nonzero eigenspace projection of u^0 and the time variable underwent a change from t to $\frac{2}{\pi}\arctan(t)$.

Another application of Theorem 6.2 presented in [10] is that, given C^∞ functions $f_0, f_1, \dots, f_p$ on a compact Riemannian manifold (M, g) of dimension n such that for all $(\theta_0, \theta_1, \dots, \theta_p) \in \mathbb{R}^{p+1} \setminus \{0\}$ the vanishing order of $\theta_0 f_0 + \dots + \theta_p f_p$ is at most γ on M, the map

$$\mathbb{RP}^p \ni [\theta_0 : \dots : \theta_p] \mapsto \partial\{\theta_0 f_0 + \dots + \theta_p f_p < 0\}$$

into the space of modulo 2 flat $(n-1)$-cycles in M is an admissible p-sweepout (see [10] for the definition). This was proposed by Marques and Neves in [74], where they settled a conjecture of Yau on the existence of infinitely many minimal surfaces in the setting of positive Ricci curvature. This proof uses an argument involving the growth in p of the min-max width $\omega_p(M)$ which is defined in terms of admissible p-sweepouts and is a nonlinear version of the spectrum of the Laplacian (cf. Gromov [41]). In fact, by Guth [43], one has

$$\omega_p(M) \le C\, p^{1/n},$$

which is reproved in [10, Theorem 3] as a corollary of Theorem 6.2.

Let us finish this section with a few consequences of the results for radicals, i.e. Theorem 4.1, due to Ghisi and Gobbino [37] and the co-area formula. These facts play an essential role in the proof of the existence of BV roots, Theorem 4.11, in [85]. They concern the behavior of differentiable functions near their zero set and the level sets of their sign; they are interesting in their own right. Note that there is no assumption on the vanishing order.

Theorem 6.6 ([85, Theorem 3.5])**.** *Let $k \in \mathbb{N}_{\ge 1}$, $\alpha \in (0,1]$, and set $s = k+\alpha$. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded Lipschitz domain and $f \in C^{k,\alpha}(\overline{\Omega}, \mathbb{R}^{\ell+1})$, $f \not\equiv 0$. Then there is a constant $C = C(n, \ell, k, \alpha, \Omega)$ such that for all $0 < \epsilon \le 1$ and all small $\delta > 0$ we have*

$$\big|\big\{y \in (0,\delta) : y^{1/s}\, \mathcal{H}^{n-1}(|f|^{-1}(y)) \ge \epsilon^{-1} C\, \|f\|^{1/s}_{C^{k,\alpha}(\overline{\Omega})}\big\}\big| \le \epsilon\delta.$$

In particular, if $A \subseteq [0,\infty)$ has full measure near 0, i.e., $|A \cap [0,\epsilon)| = \epsilon$ for some $\epsilon > 0$, then there is a sequence $A \ni y_j \to 0$ with

$$\sup_j \Big(y_j^{1/s}\, \mathcal{H}^{n-1}(|f|^{-1}(y_j))\Big) \le C\, \|f\|^{1/s}_{C^{k,\alpha}(\overline{\Omega})}.$$

That means that the $\mathcal{H}^{n-1}$-measures of the level sets $\{|f| = y_j\}$ do not grow faster than $y_j^{-1/s}$ as $y_j \to 0$.

The *sign* of $f : \Omega \to \mathbb{R}^{\ell+1}$ is the map $\operatorname{sgn}(f) : \Omega \setminus f^{-1}(0) \to \mathbb{S}^\ell$ given by

$$\operatorname{sgn}(f) := \frac{f}{|f|}.$$

Theorem 6.7 ([85, Theorem 3.3])**.** *Let $k \in \mathbb{N}_{\ge 1}$, $\alpha \in (0,1]$, and set $s = k+\alpha$. Let $\Omega \subseteq \mathbb{R}^n$ be a bounded Lipschitz domain and $f \in C^{k,\alpha}(\overline{\Omega}, \mathbb{R}^{\ell+1})$, where $n \ge \ell \ge 1$. Then there is a constant $C = C(n, \ell, k, \alpha, \Omega)$ such that for each small $\epsilon > 0$*

$$\mathcal{H}^\ell\Big(\Big\{y \in \mathbb{S}^\ell : \int_{\operatorname{sgn}(f)^{-1}(y)} |f|^{\ell/s}\, d\mathcal{H}^{n-\ell} \ge \epsilon^{-1}\, C\, \|f\|^{\ell/s}_{C^{k,\alpha}(\overline{\Omega})}\Big\}\Big) \le \epsilon. \tag{46}$$

In particular, we have ℓ/s-integrability with respect to $\mathcal{H}^{n-\ell}$ of f along almost every level set of the sign of f:

$$\int_{\operatorname{sgn}(f)^{-1}(y)} |f|^{\ell/s}\, d\mathcal{H}^{n-\ell} < \infty, \qquad \text{for } \mathcal{H}^{\ell}\text{-a.e. } y \in \mathbb{S}^{\ell}.$$

Moreover, for every relatively compact open $K \Subset \Omega \setminus f^{-1}(0)$,

$$\mathcal{H}^{n-\ell}\big(K \cap \operatorname{sgn}(f)^{-1}(y)\big) < \infty, \qquad \text{for } \mathcal{H}^{\ell}\text{-a.e. } y \in \mathbb{S}^{\ell}.$$

6.2. **Extension to the optimal transport between algebraic hypersurfaces.** A recent paper [5], by Antonini, Cavalletti, and Lerario, contains an interesting reinterpretation and extension of the regularity problem of the roots, Section 4, to the study of the Wasserstein distance on the space of d-degree hypersurfaces in $\mathbb{CP}^n$.

Let (M, g) be a smooth compact Riemannian manifold and let $\mathscr{P}(M)$ denote the space of Borel probability measures on M. Fix $q \geq 1$. For any $\mu_0, \mu_1 \in \mathscr{P}(M)$, the *$q$-Wasserstein distance* is defined by

$$W_q(\mu_0, \mu_1) := \left(\inf_{\xi \in \Pi(\mu_0,\mu_1)} \int_{M \times M} \mathsf{d}_g(x,y)^q \, \xi(dxdy) \right)^{\frac{1}{q}},$$

where d_g is the geodesic distance on M and $\Pi(\mu_0, \mu_1)$ is the set of transport plans between μ_0 and μ_1, i.e., the set of probability measures ξ on $M \times M$ with marginals μ_0 and μ_1 (the push-forward measures of ξ with respect to the natural projections). Intuitively, this distance minimizes the transport cost from μ_0 to μ_1 over all such plans.

Denote by $H_{n,d}$ the space of complex homogeneous polynomials of degree d in $n+1$ variables. Let $P_{n,d} := \mathbb{P}(H_{n,d})$ be the projectivization. Note that $P_{n,d}$ can be identified with $\mathbb{CP}^N, N = \binom{n+d}{d} - 1$. Let Z_P denote the zero set of $P \in H_{n,d}$. This set depends only on the class of P in $P_{n,d}$. Therefore, to simplify the notation, we will consider the zero sets Z_P of $P \in P_{n,d}$, abusing the terminology slightly.

Let us associate to each $P \in P_{n,d}$ a measure $\mu(P) \in \mathscr{P}(\mathbb{CP}^n)$ as follows. If Z_P is nonsingular, then we take the restriction to Z_P of the Hausdorff measure $\mathcal{H}^{2n-2}$ (with respect to the Fubini–Study metric of $\mathbb{CP}^n$), normalized:

$$\mu(P) := \frac{1}{\operatorname{vol}(Z_P)} \, \mathcal{H}^{2n-2} \llcorner Z_P,$$

where $\operatorname{vol}(Z_P) := (\mathcal{H}^{2n-2} \llcorner Z_P)(Z_P)$. (Note that all nonsingular hypersurfaces of degree d have the same volume $\operatorname{vol}(Z_P) = d \operatorname{vol}(\mathbb{CP}^{n-1})$ so that the normalization factor does not depend on P.)

Denote by $\Delta_{n,d} \subseteq P_{n,d}$ the discriminant locus, i.e., the set of polynomials $P \in P_{n,d}$ such that Z_P is singular. It is a hypersurface in $P_{n,d} = \mathbb{CP}^N$ given by

$$\Delta_{n,d} := \left\{ P \in P_{n,d} : \exists z \in \mathbb{C}^{n+1} \setminus \{0\},\ P(z) = \frac{\partial P}{\partial z_0}(z) = \cdots = \frac{\partial P}{\partial z_n}(z) = 0 \right\},$$

Using integral geometry (a Cauchy–Crofton kind of argument) one can extend the definition of $\mu(P)$ to singular hypersurfaces, that is for $P \in \Delta_{n,d}$, as follows, see [5, Theorem 1.1],

$$\int_{\mathbb{CP}^n} f \, d\mu(P) := \frac{1}{d} \int_{\mathbb{G}(1,n)} \left(\sum_{z \in Z_P \cap \ell} m_z(P|_\ell) f(z) \right) \operatorname{vol}_{\mathbb{G}(1,n)}(d\ell), \quad f \in C^0(\mathbb{CP}^n),$$

where $\mathbb{G}(1,n)$ stands for the set of projective lines in projective n-space, which is equivalent to the Grassmannian $\mathrm{Gr}(2,n+1)$, and $m_z(P|_\ell)$ is the multiplicity of z as a zero of $P|_\ell$. Then for every $q \geq 1$, the map $\mu : P_{n,d} \to \mathscr{P}_q(\mathbb{CP}^n)$ is continuous and injective. (Here by $\mathscr{P}_q(\mathbb{CP}^n)$ we mean $\mathscr{P}(\mathbb{CP}^n)$ with the Wasserstein distance W_q.)

Example 6.8 ($n = 1$)**.** For simplicity, we prefer to work with a polynomial of one complex variable $P \in \mathbb{C}[z]$. The passage to homogeneous polynomials of two complex variables z_0, z_1 is classical be means of homogenization, see e.g. [5, Remark 3.2].

Consider the space of monic polynomials of degree d in one complex variable, $P_a(z) = z^d + a_1 z^{d-1} + \cdots + a_{d-1} z + a_d$, $a = (a_1, \dots, a_d) \in \mathbb{C}^d$. To each polynomial $P = P_a \in \mathbb{C}^d$ one associates a probability measure $\mu(P) \in \mathscr{P}(\mathbb{C})$ defined by

$$\mu(P) := \frac{1}{d} \sum_{P(z)=0} m_z(P) \cdot \delta_z,$$

where $m_z(P) \in \mathbb{N}$ denotes the multiplicity of z as a zero of P and δ_z is the Dirac measure. This defines a map

$$\mu : \mathbb{C}^d \to \mathscr{P}(\mathbb{C}).$$

Geometrically, the image of μ can be identified with the quotient of $\mathbb{C}^d$ under the action of the symmetric group S_d. The q-th Wasserstein metric on $\mathscr{P}(\mathbb{C})$ is given by

$$W_q([x],[y]) = \Big(\frac{1}{d} \min_{\sigma \in \mathrm{S}_d} \sum_{j=1}^{d} |x_j - y_{\sigma(j)}|^q \Big)^{1/q},$$

where $[x] = [x_1, \dots, x_d]$ and $[y] = [y_1, \dots, y_d]$ are unordered d-tuples. It is geodesically convex and coincides with the restriction of the q-Wasserstein metric of $\mathscr{P}_q(\mathbb{C})$.

Contrary to the case $n = 1$, in general, the image of μ is not geodesically convex for the W_q-metric, that is the geodesic joining $\mu(P_0)$ to $\mu(P_1)$ in $\mathscr{P}_q(\mathbb{CP}^n)$ will in general not stay on $\mu(P_{n,d})$. To overcome this issue it is proposed in [5] to consider the inner W_q-distance, adopting the approach of J.-D. Benamou and Y. Brenier. For this one needs the notion of a q-absolutely continuous curve in this setup. A curve $\mu_t : I \to \mathscr{P}_q(\mathbb{CP}^n)$ belongs to $AC^q(I, \mathscr{P}_q(\mathbb{CP}^n))$ and is called *q-absolutely continuous* if there is $f \in L^q(I)$ such that

$$W_q(\mu_s, \mu_t) \leq \int_s^t f(\tau)\, d\tau, \quad \text{for all } s \leq t.$$

In that case, for $\mathcal{L}^1$-a.e. $t \in I$ the limit $\lim_{h\to 0} W_q(\mu_{t+h}, \mu_t)/|h|$ exists, is denoted by $|\dot\mu_t|_q$, and called *metric speed*. Then the *inner W_q-distance on $P_{n,d}$* is defined by

$$W_q^{\mathrm{in}}(P_0, P_1) := \inf_{\mu_t := \mu(\gamma(t))} \Big(\int_I |\dot\mu_t|_q^q \, dt \Big)^{\frac{1}{q}},$$

where the infimum is taken over all q-absolutely continuous curves $\gamma : I = [0,1] \to P_{n,d}$ joining P_0 and P_1.

It is shown in Section 5.2 of [5], see Theorem 5.13 and Remark 6.10 therein, that the metric space $(P_{n,d}, W_q^{\mathrm{in}})$ is complete and geodesic. Moreover, away of the discriminant locus $\Delta_{n,d}$ the W_q^{in} distance and the Fubini–Study distance d_{FS} on $P_{n,d}$ defined via the identification $P_{n,d} = \mathbb{CP}^N$ are related as follows.

Theorem 6.9. *Let $\epsilon > 0$ and denote $P_{n,d}(\epsilon) := \{p \in P_{n,d} : \mathrm{d_{FS}}(p, \Delta_{n,d}) \geq \epsilon\}$. Then the identity map*

$$\mathrm{id} : (P_{n,d}(\epsilon), \mathrm{d_{FS}}) \to (P_{n,d}(\epsilon), W_q^{\mathrm{in}})$$

is Lipschitz. In particular, $P_{n,d}(\epsilon)$ is compact in the W_q^{in}–topology.

In the proof, continuous semialgebraic curves are used. By Puiseux's theorem, they can be reparametrized, near their singular points, to get C^m regularity, for arbitrary finite m.

Another important question is the compactness of $(P_{n,d}, W_q^{\mathrm{in}})$. This problem was only partially solved in [5].

Theorem 6.10 ([5, Theorem 6.11])**.** *There is $e(n) \in \mathbb{N}$ such that, for every $1 \leq q < e(n)d/(e(n)d-1)$, the complete and separable geodesic space $(P_{n,d}, W_q^{\mathrm{in}})$ is compact.*

Theorem 4.2 is used in the proof of the above theorem to provide q-absolutely continuous curves as follows. This is fairly straightforward for $n = 1$ because, any curve of polynomials

$$P_a(t)(z) = z^d + \sum_{j=1}^{d} a_j(t) z^{d-j}, \quad t \in [0,1],$$

with $a_j \in C^{d-1,1}([0,1])$, defines, after homogenization, an element of $AC^q([0,1]; \mathscr{P}_q(\mathbb{CP}^1))$ for every $1 \leq q < d/(d-1)$. Thus one may take $e(1) = 1$.

Thanks to the uniformity of the bound of Theorem 4.2, the general case can be reduced to the case $n = 1$ by restriction to all affine lines parameterized by $\mathbb{G}(1,n)$ and then integration over this Grassmannian, provided we can avoid the lines entirely included in Z_P. This is always possible for generic P if $d > 2n - 3$, [5, Theorem 6.3], but not in general. Therefore it is proposed in [5] to use the integration over the group of unitary transformations of a rational curve of sufficiently high degree (instead of a line). It follows from a recent result of B. Lehmann, E. Riedl and S. Tanimoto that for a given curve of polynomials $P_t \in P_{n,d}$ there is a rational curve of degree $e(n)$, given by an explicit formula, such that none of its unitary transformations can be entirely included in Z_{P_t}. For all details of the proof we refer the reader to [5].

Open Problem 6.11. *For which $q \geq 1$ is the space $(P_{n,d}, W_q^{\mathrm{in}})$ compact? In particular, is it always compact for $q = 2$?*

7. Appendix A. Function spaces

7.1. **Hölder spaces.** Let $\Omega \subseteq \mathbb{R}^n$ be open. We denote by $C^0(\Omega)$ the space of continuous complex valued functions on Ω. For $k \in \mathbb{N}_{\geq 1} \cup \{\infty\}$, $C^k(\Omega)$ is the space of k-times continuously differentiable functions,

$$C^k(\Omega) = \{f \in \mathbb{C}^\Omega : \partial^\alpha f \in C^0(\Omega) \text{ for } 0 \leq |\alpha| \leq k\}, \quad k \in \mathbb{N},$$

and

$$C^\infty(\Omega) = \bigcap_{k \in \mathbb{N}} C^k(\Omega).$$

The space of real analytic functions on Ω is denoted by $C^\omega(\Omega)$. If $\Omega \subseteq \mathbb{C}^n$, then $\mathcal{H}(\Omega)$ is the space of holomorphic functions on Ω.

If the open set $\Omega \subseteq \mathbb{R}^n$ is bounded, we write $C^k(\overline{\Omega})$ for the space of all $f \in C^k(\Omega)$ such that $\partial^\alpha f$, $0 \le |\alpha| \le k$, has a continuous extension to the closure $\overline{\Omega}$. Then $C^k(\overline{\Omega})$ endowed with the norm

$$\|f\|_{C^k(\overline{\Omega})} := \sup_{|\alpha|\le k} \sup_{x\in\Omega} |\partial^\alpha f(x)|$$

is a Banach space. The spaces $C^k(\Omega)$, $C^\infty(\Omega)$, $C^\omega(\Omega)$, and $\mathcal{H}(\Omega)$ carry their natural locally convex topologies, where Ω is not necessarily bounded; note that $C^k(\Omega)$, $C^\infty(\Omega)$, and $\mathcal{H}(\Omega)$ are Fréchet spaces.

Let $\alpha \in (0,1]$ and let $\Omega \subseteq \mathbb{R}^n$ be open and bounded. A function is *α-Hölder continuous in* Ω if

$$\operatorname{Hölld}_{\alpha,\Omega}(f) := \sup_{x,y\in\Omega,\, x\ne y} \frac{|f(x)-f(y)|}{|x-y|^\alpha} < \infty.$$

It is *Lipschitz continuous in* Ω if it is 1-Hölder continuous in Ω; in that case, we write

$$\operatorname{Lip}_\Omega(f) := \operatorname{Hölld}_{1,\Omega}(f) < \infty.$$

For $k \in \mathbb{N}$ and $\alpha \in (0,1]$, we have the scale of *Hölder(–Lipschitz) spaces*

$$C^{k,\alpha}(\overline{\Omega}) := \{f \in C^k(\overline{\Omega}) : \partial^\beta f \text{ is } \alpha\text{-Hölder continuous for } |\beta| = k\}$$

and we equip them with the norm

$$\|f\|_{C^{k,\alpha}(\overline{\Omega})} := \|f\|_{C^k(\overline{\Omega})} + \sup_{|\beta|=k} \operatorname{Hölld}_{\alpha,\Omega}(\partial^\beta f).$$

Then $C^{k,\alpha}(\overline{\Omega})$ is a Banach space. We write $C^{k,\alpha}(\Omega)$ (where Ω is not necessarily bounded) for the space of all functions $f \in C^k(\Omega)$ such that $\|f|_{\Omega'}\|_{C^{k,\alpha}(\overline{\Omega}')} < \infty$ for all relatively compact $\Omega' \Subset \Omega$ with its natural Fréchet topology.

7.2. **Lebesgue spaces and weak Lebesgue spaces.** Let $\Omega \subseteq \mathbb{R}^n$ be open and $1 \le p \le \infty$. The *Lebesgue space* $L^p(\Omega)$ is the space of measurable complex valued functions $f : \Omega \to \mathbb{C}$ that are p-integrable with respect to the n-dimensional Lebesgue measure $\mathcal{L}^n$; actually, the elements of $L^p(\Omega)$ are equivalence classes of functions that coincide $\mathcal{L}^n$-almost everywhere. If not stated otherwise, "measurable" always means "Lebesgue measurable". The L^p-norms make the Lebesgue spaces to Banach spaces:

$$\|f\|_{L^p(\Omega)} := \Big(\int_\Omega |f(x)|^p dx\Big)^{1/p}, \quad 1 \le p < \infty,$$
$$\|f\|_{L^\infty(\Omega)} := \operatorname{ess\,sup}_{x\in\Omega} |f(x)|, \quad p = \infty.$$

We denote by $L^p_{\mathrm{loc}}(\Omega)$ the space of measurable functions $f : \Omega \to \mathbb{C}$ such that $\|f\|_{L^p(K)} < \infty$ for each compact subset $K \subseteq \Omega$.

Let $\Omega \subseteq \mathbb{R}^n$ be open and bounded. For $1 \le p < \infty$, the *weak L^p-space* is the space of all measurable functions $f : \Omega \to \mathbb{C}$ such that

$$\|f\|_{p,w,\Omega} := \sup_{r>0} \big(r \cdot \mathcal{L}^n(\{x \in \Omega : |f(x)| > r\})^{1/p}\big) < \infty;$$

again, we identify functions that coincide $\mathcal{L}^n$-almost everywhere. Note that $\|\cdot\|_{p,w,\Omega}$ is only a quasinorm: the triangle inequality fails, but for $f_j \in L^p_w(\Omega)$ one still has

$$\Big\|\sum_{j=1}^m f_j\Big\|_{p,w,\Omega} \le m \sum_{j=1}^m \|f_j\|_{p,w,\Omega}.$$

The spaces $L^p_w(\Omega)$ are complete with respect to the quasinorms $\|\cdot\|_{p,w,\Omega}$, i.e., they are quasi-Banach spaces. If $p > 1$, then there exists a norm equivalent to $\|\cdot\|_{p,w,\Omega}$ which makes $L^p_w(\Omega)$ into a Banach space. For $1 \le q < p < \infty$, we have

$$\|f\|_{q,w,\Omega} \le \|f\|_{L^q(\Omega)} \le \Big(\frac{p}{p-q}\Big)^{1/q} \mathcal{L}^n(\Omega)^{1/q-1/p}\, \|f\|_{p,w,\Omega}$$

and thus there are continuous strict inclusions $L^p_w(\Omega) \subseteq L^q(\Omega) \subseteq L^q_w(\Omega)$. The L^p_w-quasinorm is σ-subadditive (i.e., $\|f\|^p_{p,w,\Omega} \le \sum_j \|f\|^p_{p,w,\Omega_j}$ for countable open covers $\Omega = \bigcup_j \Omega_j$) but not σ-additive.

7.3. **Sobolev spaces.** Let $\Omega \subseteq \mathbb{R}^n$ be open. For $k \in \mathbb{N}_{\ge 1}$ and $1 \le p \le \infty$, we consider the *Sobolev space*

$$W^{k,p}(\Omega) := \{f \in L^p(\Omega) : \partial^\alpha f \in L^p(\Omega) \text{ for } |\alpha| \le k\},$$

where $\partial^\alpha f$ are distributional derivatives, with the norm

$$\|f\|_{W^{k,p}(\Omega)} := \sum_{|\alpha|\le k} \|\partial^\alpha f\|_{L^p(\Omega)}.$$

Then $W^{k,p}(\Omega)$ are Banach spaces. We denote by $W^{k,p}_{\mathrm{loc}}(\Omega)$ the space of functions $f \in L^p_{\mathrm{loc}}(\Omega)$ such that $\partial^\alpha f \in L^p_{\mathrm{loc}}(\Omega)$ for $|\alpha| \le k$.

Let $k = 1$ and suppose $f \in L^p(\Omega)$. Then $f \in W^{1,p}(\Omega)$ if and only if f has a representative $\overline{f}$ that is absolutely continuous on almost all line segments in Ω parallel to the coordinate axes and whose classical derivatives $\partial_i \overline{f}$ belong to $L^p(\Omega)$. In particular, for a bounded interval $I \subseteq \mathbb{R}$, $W^{1,1}(I)$ can be identified with the space $AC(I)$ of absolutely continuous functions on I. We recall that, by definition, $f \in AC(I)$ if for every $\epsilon > 0$ there exists $\delta > 0$ such that $\sum_{j=1}^k |f(b_j) - f(a_j)| \le \epsilon$ for every finite number of disjoint intervals (a_j, b_j), $j = 1, \dots, k$, satisfying $[a_j, b_j] \subseteq I$ and $\sum_{j=1}^k (b_j - a_j) \le \delta$.

In the case that $p = \infty$ and that Ω is a bounded Lipschitz domain, $f \in W^{1,\infty}(\Omega)$ if and only if f has a representative that is Lipschitz continuous on Ω.

7.4. **Functions of bounded variation.** Let $\Omega \subseteq \mathbb{R}^n$ be open. A real valued $f \in L^1(\Omega)$ is a *function of bounded variation in* Ω if the distributional derivative of f is representable by a finite Radon measure in Ω, i.e., for all C^∞ functions $\varphi : \Omega \to \mathbb{R}$ with compact support in Ω and all $1 \le i \le n$,

$$\int_\Omega f(x)\partial_i\varphi(x)\, dx = -\int_\Omega \varphi\, dD_i f,$$

for some $\mathbb{R}^n$-valued Radon measure $Df = (D_1 f, \dots, D_n f)$ on Ω. By definition, a complex valued function is of bounded variation in Ω if its real and imaginary parts are. The space $BV(\Omega)$ of all functions of bounded variation in Ω equipped with the norm

$$\|f\|_{BV(\Omega)} := \|f\|_{L^1(\Omega)} + |Df|(\Omega),$$

where $|Df|(\Omega)$ is the total variation measure, is a Banach space.

The Sobolev space $W^{1,1}(\Omega)$ is strictly contained in $BV(\Omega)$ (for $f \in W^{1,1}(\Omega)$ we have $Df = \nabla f\, \mathcal{L}^n$).

Functions of bounded variation may have discontinuities. Let $f \in BV(\Omega)$. Then the Lebesgue decomposition with respect to $\mathcal{L}^n$ induces a decomposition

$$Df = D^a f + D^j f + D^c f,$$

where $D^a f = \nabla f\, \mathcal{L}^n$ is the *absolutely continuous part*, $D^j f$ is the *jump part* (given by integration of the jump height against the restriction of $\mathcal{H}^{n-1}$ to the set of approximate jump points J_f, i.e., $D^j f = ((f^+ - f^-) \otimes \nu_f)\,(\mathcal{H}^{n-1} \llcorner J_f)$, where ν_f is the direction of the jump), and $D^c f$ is the *Cantor part*.

The closed subspace $SBV(\Omega) := \{f \in BV(\Omega) : D^c f = 0\}$ of $BV(\Omega)$ is called the space of *special functions of bounded variation*. We have continuous strict inclusions $W^{1,1}(\Omega) \subseteq SBV(\Omega) \subseteq BV(\Omega)$.

8. Appendix B. The space of hyperbolic polynomials

The fact that the regularity of the roots of hyperbolic polynomials is essentially better than that of the roots of general polynomials is related to the geometry of the space $\mathrm{Hyp}(d)$ of monic hyperbolic polynomials of degree d and the constraints a sufficiently differentiable curve of hyperbolic polynomials is subdued to at the boundary of $\mathrm{Hyp}(d)$. In this section, we collect some interesting topological and geometric properties of $\mathrm{Hyp}(d)$.

The following results are due to Kostov [54]; see also Arnol'd [6] and Givental' [38]. Let $p = (p_1, \dots, p_d)$ be a probability vector, i.e., $p_i > 0$ and $p_1 + \dots + p_d = 1$. The *Vandermonde mapping* $W = W(p) : \mathbb{R}^d \to \mathbb{R}^d$ is defined by

$$W_j(x) := p_1 x_1^j + \dots + p_d x_d^j, \quad 1 \le j \le d.$$

If $p_i = 1/d$ for $1 \le j \le d$, then W_j, $1 \le j \le d$, are the Newton polynomials (up to a constant factor) which generate the algebra of symmetric polynomials on $\mathbb{R}^d$ and consequently the image $W(\mathbb{R}^d)$ is isomorphic to $\mathrm{Hyp}(d)$ in this case.

The set $K := \{x \in \mathbb{R}^d : x_1 \le x_2 \le \dots \le x_d\}$ is a fundamental domain for the action of the symmetric group S_d on $\mathbb{R}^d$ by permuting the coordinates. Every set of the form $\{x \in K : W_j(x) = c_j,\ 1 \le j \le k\}$ is called a $(d-k)$-*dimensional Vandermonde manifold*.

Theorem 8.1 ([54])**.** *Every Vandermonde manifold is either contractible or empty. The map $W : K \to W(K)$ is a homeomorphism. Every set of the form $\{y \in W(K) : y_j = c_j,\ 1 \le j \le k\}$ is either contractible or empty.*

As a consequence, $W^k := (W_1, \dots, W_k)$, $k \le d$, is a homeomorphism of the closure of every k-dimensional stratum of the polyhedron K onto its image; indeed, the restriction of W^k to such a stratum is the k-dimensional Vandermonde mapping for some probability vector.

Let $K_0 := \{x \in K : W_1(x) = 0,\ W_2(x) \le 1\}$.

Theorem 8.2 ([54])**.** *The sets $W(K_0)$ and $W^k(K_0)$, $1 \le k \le d-1$, are quasiconvex, i.e., there is a constant $C > 0$ such that any two points can be joint by a piecewise smooth curve in the set of length not exceeding C times the Euclidean distance of the two points.*

In particular, the space of monic hyperbolic polynomials of degree d in Tschirnhausen form with $|\tilde{a}_2| \le 1$ is quasiconvex. We refer to [54] for additional information on the stratification of $W(K)$.

The semialgebraic set $\mathrm{Hyp}(d) \subseteq \mathbb{R}^d$ can be described by explicit inequalities. Let $s_j = x_1^j + \dots + x_d^j$, $j \ge 0$, be the Newton polynomials and consider the Bezoutiant

matrix

$$B := \begin{pmatrix} s_0 & s_1 & \cdots & s_{d-1} \\ s_1 & s_2 & \cdots & s_d \\ \vdots & \vdots & \ddots & \vdots \\ s_{d-1} & s_d & \cdots & s_{2d-2} \end{pmatrix}.$$

We have $B = \tilde{B}(\sigma_1, \dots, \sigma_d)$ for a unique matrix valued polynomial $\tilde{B}$, where the σ_i are the elementary symmetric functions. The following result is Hermite's [47] and Sylvester's [107] version of Sturm's theorem; the indicated references give a modern account and a far reaching generalization.

Theorem 8.3 ([89, 90])**.** *The space* $\mathrm{Hyp}(d)$ *of monic hyperbolic polynomials of degree* d *can be identified with* $\{y \in \mathbb{R}^d : \tilde{B}(y) \geq 0\}$*. The rank of* $\tilde{B}$ *equals the number of distinct roots and its signature equals the number of distinct real roots.*

Since a real symmetric matrix is positive semidefinite if and only if the determinants of the principal minors are nonnegative, this description yields explicit inequalities that determine $\mathrm{Hyp}(d)$.

Recall that $\mathrm{Hyp}_T(d) = \mathrm{Hyp}(d) \cap \{y_1 = 0\}$ is the space of monic hyperbolic polynomials of degree d in Tschirnhausen form. Then $\mathrm{Hyp}_T(1)$ is the origin, $\mathrm{Hyp}_T(2)$ is $(-\infty, 0]$, $\mathrm{Hyp}_T(3)$ is the closure of the inside of a cusp, and $\mathrm{Hyp}_T(4)$ is the closure of the inside of the swallowtail.

In [75] Meguerditchian studies the "escape" from the set $\mathrm{Hyp}(d)$: roughly speaking, for a polynomial P in $\mathrm{Hyp}(d)$ an integer s_P, which depends only on the multiplicity vector of P, is introduced such that one can decide whether $P + Q$ belongs to $\mathrm{Hyp}(d)$ or not in terms of how the degree of Q relates to s_P.

By Nuij [79], for $P \in \mathrm{Hyp}(d)$ also $P + sP' \in \mathrm{Hyp}(d)$, for all $s \in \mathbb{R}$. A generalization of this result is due to Kurdyka and Păunescu [65].

Acknowledgements. Armin Rainer was supported by the Austrian Science Fund (FWF) Project P 32905-N.

References

[1] S. S. Abhyankar, *On the ramification of algebraic functions*, Amer. J. Math., **77**, (1955), 575–592.

[2] D. Alekseevsky, A. Kriegl, M. Losik, and P. W. Michor, *Choosing roots of polynomials smoothly*, Israel J. Math. **105** (1998), 203–233.

[3] ———, *Lifting smooth curves over invariants for representations of compact Lie groups*, Transform. Groups **5** (2000), no. 2, 103–110.

[4] F. J. Almgren, Jr., *Almgren's big regularity paper*, World Scientific Monograph Series in Mathematics, vol. 1, World Scientific Publishing Co. Inc., River Edge, NJ, 2000. Q-valued functions minimizing Dirichlet's integral and the regularity of area-minimizing rectifiable currents up to codimension 2, With a preface by Jean E. Taylor and Vladimir Scheffer.

[5] P. Antonini, F. Cavalletti, A. Lerario, *Optimal transport between algebraic hypersurfaces*, preprint, arXiv:2212.10274

[6] V. I. Arnol′d, *Hyperbolic polynomials and Vandermonde mappings*, Funktsional. Anal. i Prilozhen. **20** (1986), no. 2, 52–53.

[7] M. Bačák and J. M. Borwein, *On difference convexity of locally Lipschitz functions*, Optimization **60** (2011), no. 8-9, 961–978.

[8] C. Bär, *Zero sets of solutions to semilinear elliptic systems of first order*, Invent. Math. **138** (1999), no. 1, 183–202.

[9] H. H. Bauschke, O. Güler, A. S. Lewis, and H. S. Sendov, *Hyperbolic polynomials and convex analysis*, Canad. J. Math. **53** (2001), no. 3, 470–488.

[10] T. Beck, S. Becker-Kahn, and B. Hanin, *Nodal sets of smooth functions with finite vanishing order and p-sweepouts*, Calc. Var. Partial Differential Equations **57** (2018), no. 5, 1–13.
[11] R. Bhatia, *Matrix analysis*, Graduate Texts in Mathematics, vol. 169, Springer-Verlag, New York, 1997.
[12] R. Bhatia, C. Davis, and A. McIntosh, *Perturbation of spectral subspaces and solution of linear operator equations*, Linear Algebra Appl. **52/53** (1983), 45–67.
[13] E. Bierstone, *Lifting isotopies from orbit spaces*, Topology **14** (1975), no. 3, 245–252.
[14] E. Bierstone, P. Milman, *Arc-analytic functions*, Invent. Math. **101** (1990), no. 1, 411–424.
[15] ———, *Resolution of singularities in Denjoy-Carleman classes*, Selecta Math. (N.S.) **10** (2004), no. 1, 1–28.
[16] J. Bochnak, M. Coste, and M.-F. Roy, *Real algebraic geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 36, Springer-Verlag, Berlin, 1998, Translated from the 1987 French original, Revised by the authors.
[17] J.-M. Bony, *Sommes de carrés de fonctions dérivables*, Bull. Soc. Math. France **133** (2005), no. 4, 619–639.
[18] J.-M. Bony, F. Broglia, F. Colombini, and L. Pernazza, *Nonnegative functions as squares or sums of squares*, J. Funct. Anal. **232** (2006), no. 1, 137–147.
[19] J.-M. Bony, F. Colombini, and L. Pernazza, *On square roots of class C^m of nonnegative functions of one variable*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **9** (2010), no. 3, 635–644.
[20] J. Borcea and P. Brändén, *Multivariate Pólya-Schur classification problems in the Weyl algebra*, Proceedings of the London Mathematical Society **101** (2010), no. 1, 73–104.
[21] M. D. Bronshtein, *Smoothness of roots of polynomials depending on parameters*, Sibirsk. Mat. Zh. **20** (1979), no. 3, 493–501, 690, English transl. in Siberian Math. J. **20** (1980), 347–352.
[22] ———, *The Cauchy problem for hyperbolic operators with characteristics of variable multiplicity*, Trudy Moskov. Mat. Obshch. **41** (1980), 83–99, English transl. in Trans. Moscow Math. Soc. **1**(1982), 87–103.
[23] C. Chevalley, *Invariants of finite groups generated by reflections.*, Am. J. Math. **77** (1955), 778–782 (English).
[24] P. M. Cohn, Puiseux's theorem revisited, *J. Pure Appl. Algebra*, **31**, (1984) 1–4; Corrigendum: *J. Pure Appl. Algebra*, **52**, (1988), 197–198.
[25] F. Colombini, E. Jannelli, and S. Spagnolo, *Well-posedness in the Gevrey classes of the Cauchy problem for a nonstrictly hyperbolic equation with coefficients depending on time*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **10** (1983), no. 2, 291–312.
[26] F. Colombini and N. Lerner, *Une procedure de Calderón-Zygmund pour le problème de la racine k-ième*, Ann. Mat. Pura Appl. (4) **182** (2003), no. 2, 231–246.
[27] F. Colombini, N. Orrù, and L. Pernazza, *On the regularity of the roots of hyperbolic polynomials*, Israel J. Math. **191** (2012), 923–944.
[28] J. Dadok, *Polar coordinates induced by actions of compact Lie groups*, Trans. Amer. Math. Soc. **288** (1985), no. 1, 125–137.
[29] J. Dadok and V. Kac, *Polar representations*, J. Algebra **92** (1985), no. 2, 504–524.
[30] C. De Lellis and E. N. Spadaro, *Q-valued functions revisited*, Mem. Amer. Math. Soc. **211** (2011), no. 991, vi+79.
[31] H. Derksen and G. Kemper, *Computational invariant theory*, Invariant Theory and Algebraic Transformation Groups, I, Springer-Verlag, Berlin, 2002, Encyclopaedia of Mathematical Sciences, 130.
[32] H. Donnelly and C. Fefferman, *Nodal sets of eigenfunctions on Riemannian manifolds*, Invent. Math. **93** (1988), no. 1, 161–183.
[33] R. M. Dudley, *On second derivatives of convex functions*, Math. Scand. **41** (1977), no. 1, 159–174.
[34] C. Fefferman and D. H. Phong, *On positivity of pseudo-differential operators*, Proc. Nat. Acad. Sci. U.S.A. **75** (1978), no. 10, 4673–4674.
[35] L. Gårding, *Linear hyperbolic partial differential equations with constant coefficients*, Acta Math. **85** (1951), 1–62.
[36] ———, *An inequality for hyperbolic polynomials*, J. Math. Mech. **8** (1959), 957–965.
[37] M. Ghisi and M. Gobbino, *Higher order Glaeser inequalities and optimal regularity of roots of real functions*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **12** (2013), no. 4, 1001–1021.
[38] A. B. Givental′, *Moments of random variables and the equivariant Morse lemma*, Uspekhi Mat. Nauk **42** (1987), no. 2(254), 221–222.

[39] G. Glaeser, *Racine carrée d'une fonction différentiable*, Ann. Inst. Fourier (Grenoble) **13** (1963), no. 2, 203–210.

[40] V. Grandjean, Re-parameterizing and reducing families of normal operators, *Israel Journal of Mathematics*, **230**, no. 2, (2019), 715–744.

[41] M. Gromov, *Dimension, nonlinear spectra and width*, Geometric aspects of functional analysis (1986/87), Lecture Notes in Math., vol. 1317, Springer, Berlin, 1988, pp. 132–184.

[42] L. Gurvits, *Van der Waerden/Schrijver-Valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: one theorem for all*, Electron. J. Combin. **15** (2008), no. 1, Research Paper 66, 26, With a corrigendum.

[43] L. Guth, *Minimax problems related to cup powers and Steenrod squares*, Geom. Funct. Anal. **18** (2009), no. 6, 1917–1987.

[44] P. Hartman, *On functions representable as a difference of convex functions*, Pacific J. Math. **9** (1959), no. 3, 707–713.

[45] F. R. Harvey and H. B. Lawson, Jr., *Gårding's theory of hyperbolic polynomials*, Comm. Pure Appl. Math. **66** (2013), no. 7, 1102–1128.

[46] J. W. Helton and V. Vinnikov, *Linear matrix inequality representation of sets*, Commun. Pure Appl. Math. **60** (2007), no. 5, 654–674 (English).

[47] C. Hermite, *Remarques sur le théorème de Sturm*, C. R. Acad. Sci. Paris **36** (1853), 52-54.

[48] J.-B. Hiriart-Urruty, *Generalized differentiability, duality and optimization for problems dealing with differences of convex functions*, Convexity and duality in optimization (Groningen, 1984), Lecture Notes in Econom. and Math. Systems, vol. 256, Springer, Berlin, 1985, pp. 37–70.

[49] H. Hironaka, *Resolution of singularities of an algebraic variety over a field of characteristic zero. I*, Ann. of Math. (2) **79** (1964), 109—203.

[50] D. Jerison and G. Lebeau, *Nodal sets of sums of eigenfunctions*, Harmonic analysis and partial differential equations (Chicago, IL, 1996), Chicago Lectures in Math., Univ. Chicago Press, Chicago, IL, 1999, pp. 223–239.

[51] H. E. W. Jung, *Darstellung der Funktionen eines algebraischen Körpers zweier unabhängiger Varänderlichen x, y in der Umbebung einer Stelle $x = a$, $y = b$*, J. Reine Angew. Math. **133** (1908), 289–314.

[52] T. Kato, *Perturbation theory for linear operators*, second ed., Grundlehren der Mathematischen Wissenschaften, vol. 132, Springer-Verlag, Berlin, 1976.

[53] S. I. Khudyaev and A. I. Vol'pert, *Analysis in classes of discontinuous functions and equations of mathematical physics*, Mechanics: Analysis, 8. Martinus Nijhoff Publishers, Dordrecht, 1985.

[54] V. P. Kostov, *On the geometric properties of Vandermonde's mapping and on the problem of moments*, Proc. Roy. Soc. Edinburgh Sect. A **112** (1989), no. 3-4, 203–211.

[55] A. Kriegl, M. Losik, and P. W. Michor, *Choosing roots of polynomials smoothly. II*, Israel J. Math. **139** (2004), 183–188.

[56] A. Kriegl, M. Losik, P. W. Michor, and A. Rainer, *Lifting smooth curves over invariants for representations of compact Lie groups. II*, J. Lie Theory **15** (2005), no. 1, 227–234.

[57] ———, *Lifting smooth curves over invariants for representations of compact Lie groups. III*, J. Lie Theory **16** (2006), no. 3, 579–600.

[58] ———, *Lifting mappings over invariants of finite groups*, Acta Math. Univ. Comenian. (N.S.) **77** (2008), no. 1, 93–122.

[59] ———, *Addendum to: "Lifting smooth curves over invariants for representations of compact Lie groups. III"[J. Lie Theory 16 (2006), no. 3, 579–600]*, J. Lie Theory **22** (2012), no. 1, 245–249.

[60] A. Kriegl and P. W. Michor, *The Convenient Setting of Global Analysis*, Mathematical surveys and monographs **53**, American Mathematical Society, Providence, RI, 1997

[61] ———, *Differentiable perturbation of unbounded operators*, Math. Ann. **327** (2003), no. 1, 191–201.

[62] A. Kriegl, P. W. Michor, and A. Rainer, *Denjoy–Carleman differentiable perturbation of polynomials and unbounded operators*, Integral Equations and Operator Theory **71** (2011), no. 3, 407–416.

[63] ———, *Many parameter Hölder perturbation of unbounded operators*, Math. Ann. **353** (2012), 519–522.

[64] K. Kurdyka, L. Păunescu, *Hyperbolic polynomials and multiparameter real-analytic perturbation theory*, Duke Math. J., **141**, no. 1, (2008), 123-149.
[65] ———, *Nuij type pencils of hyperbolic polynomials*, Canad. Math. Bull. **60** (2017), no. 3, 561–570.
[66] P. D. Lax, *Differential equations, difference equations and matrix theory*, Comm. Pure Appl. Math. **11** (1958), 175–194.
[67] A. S. Lewis, P. A. Parrilo, and M. V. Ramana, *The Lax conjecture is true.*, Proc. Am. Math. Soc. **133** (2005), no. 9, 2495–2499 (English).
[68] F.-H. Lin, *Nodal sets of solutions of elliptic and parabolic equations*, Comm. Pure Appl. Math. **44** (1991), no. 3, 287–308.
[69] S. Łojasiewicz, *Ensembles semi-analytiques*, Inst. Hautes Études Sci., Bures-sur-Yvette, 1965.
[70] M. Losik, P. W. Michor, and A. Rainer, *A generalization of Puiseux's theorem and lifting curves over invariants*, Rev. Mat. Complut. **25** (2012), no. 1, 139–155.
[71] M. Losik and A. Rainer, *Choosing roots of polynomials with symmetries smoothly*, Rev. Mat. Complut. **20** (2007), no. 2, 267–291.
[72] D. Luna, *Slices étales*, Sur les groupes algébriques, Soc. Math. France, Paris, 1973, pp. 81–105. Bull. Soc. Math. France, Paris, Mémoire 33.
[73] A. W. Marcus, D. A. Spielman, and N. Srivastava, *Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem*, Ann. of Math. (2) **182** (2015), no. 1, 327–350.
[74] F. C. Marques and A. Neves, *Existence of infinitely many minimal hypersurfaces in positive Ricci curvature*, Invent. Math. **209** (2017), no. 2, 577–616.
[75] I. Meguerditchian, *A theorem on the escape from the space of hyperbolic polynomials*, Math. Z. **211** (1992), no. 3, 449–460.
[76] P. W. Michor, *Topics in differential geometry*, Graduate Studies in Mathematics, vol. 93, American Mathematical Society, Providence, RI, 2008.
[77] C. Miller, *Exponentiation is hard to avoid*, Proc. Amer. Math. Soc. **122** (1994), no. 1, 257–259.
[78] S. Mukai, *An introduction to invariants and moduli. Transl. by W. M. Oxbury.*, vol. 81, Cambridge: Cambridge University Press, 2003 (English).
[79] W. Nuij, *A note on hyperbolic polynomials*, Math. Scand. **23** (1968), 69–72 (1969).
[80] Y. Ohya and S. Tarama, *Le problème de Cauchy à caractéristiques multiples dans la classe de Gevrey. I. Coefficients hölderiens en t*, Hyperbolic equations and related topics (Katata/Kyoto, 1984), Academic Press, Boston, MA, 1986, pp. 273–306.
[81] A. Parusiński and A. Rainer, *A new proof of Bronshtein's theorem*, J. Hyperbolic Differ. Equ. **12** (2015), no. 4, 671–688.
[82] ———, *Lifting differentiable curves from orbit spaces*, Transform. Groups **21** (2016), no. 1, 153–179.
[83] ———, *Regularity of roots of polynomials*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **16** (2016), 481–517.
[84] ———, *Optimal Sobolev regularity of roots of polynomials*, Ann. Scient. Éc. Norm. Sup. 4 e série **51** (2018), no. 5, 1343–1387.
[85] ———, *Selections of bounded variation for roots of smooth polynomials*, Sel. Math. New Ser. **26**, 13 (2020).
[86] ———, *Sobolev lifting over invariants*, SIGMA Symmetry Integrability Geom. Methods Appl. **17** (2021), Paper no. 037.
[87] A. Parusiński, G. Rond, *The Abhyankar-Jung Theorem,* Journal of Algebra **365** (2012), 29–41.
[88] ———, *Multiparameter perturbation theory of matrices and linear operators*, Trans. Amer. Math. Soc. **373** (2020) 2933-2948.
[89] C. Procesi, *Positive symmetric functions*, Adv. in Math. **29** (1978), no. 2, 219–225.
[90] C. Procesi and G. Schwarz, *Inequalities defining orbit spaces*, Invent. Math. **81** (1985), no. 3, 539–554.
[91] Q. I. Rahman and G. Schmeisser, *Analytic theory of polynomials*, London Mathematical Society Monographs. New Series, vol. 26, The Clarendon Press Oxford University Press, Oxford, 2002.
[92] A. Rainer, *Smooth roots of hyperbolic polynomials with definable coefficients*, Israel J. Math. **184** (2011), 157–182.

[93] ———, *Lifting quasianalytic mappings over invariants.*, Canad. J. Math. **64** (2012), no. 2, 409–428.

[94] ———, *Perturbation theory for normal operators*, Trans. Amer. Math. Soc. **365** (2013), no. 10, 5545–5577.

[95] ———, *Differentiable roots, eigenvalues, and eigenvectors*, Israel J. Math. **201** (2014), no. 1, 99–122.

[96] ———, *Roots of Gårding hyperbolic polynomials*, Proc. Amer. Math. Soc. **150** (2022), no. 6, 2433–2446.

[97] ———, *Quantitative tame properties of differentiable functions with controlled derivatives*, Nonlinear Analysis **237** (2023), 113372.

[98] K. Ray and J. Schmidt-Hieber, *A regularity class for the roots of nonnegative functions*, Ann. Mat. Pura Appl. (4) **196** (2017), no. 6, 2091–2103.

[99] F. Rellich, *Störungstheorie der Spektralzerlegung, I*, Math. Ann. **113** (1937), 600–619.

[100] ———, *Perturbation theory of eigenvalue problems*, Assisted by J. Berkowitz. With a preface by Jacob T. Schwartz, Gordon and Breach Science Publishers, New York, 1969.

[101] R. T. Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970.

[102] G. W. Schwarz, *Lifting smooth homotopies of orbit spaces*, Inst. Hautes Études Sci. Publ. Math. **51** (1980), 37–135.

[103] J.-P. Serre, *Groupes finis d'automorphismes d'anneaux locaux réguliers.*, Colloque d'algèbre (Paris, 1967), Secrétariat mathématique, Paris, Exp. 8, 11 p. (1968)., 1968.

[104] G. C. Shephard and J. A. Todd, *Finite unitary reflection groups.*, Can. J. Math. **6** (1954), 274–304 (English).

[105] S. Spagnolo, *On the absolute continuity of the roots of some algebraic equations*, Ann. Univ. Ferrara Sez. VII (N.S.) **45** (1999), no. suppl., 327–337 (2000), Workshop on Partial Differential Equations (Ferrara, 1999).

[106] ———, *Local and semi-global solvability for systems of non-principal type*, Comm. Partial Differential Equations **25** (2000), no. 5-6, 1115–1141.

[107] J.J. Sylvester, *On a theory of syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's function*, Philos. Trans. Roy. Soc. London **143** (1853).

[108] S. Tarama, *On the lemma of Colombini, Jannelli and Spagnolo*, Memoirs of the Faculty of Engineering, Osaka City University **41** (2000), 111–115.

[109] ———, *Note on the Bronshtein theorem concerning hyperbolic polynomials*, Sci. Math. Jpn. **63** (2006), no. 2, 247–285.

[110] C.-L. Terng, *Isoparametric submanifolds and their Coxeter groups*, J. Differential Geom. **21** (1985), no. 1, 79–107.

[111] A. Tognoli, *Sulla classificazione dei fibrati analitici reali E-principali*, Ann. Scuola Norm. Sup. Pisa (3) **23** (1969), 75–86.

[112] L. Veselý, J. Duda, and L. Zajiček, *On d.c. functions and mappings.*, Atti Semin. Mat. Fis. Univ. Modena **51** (2003), no. 1, 111–138 (English).

[113] V. Vinnikov, *Self-adjoint determinantal representations of real plane curves*, Math. Ann. **296** (1993), no. 1, 453–479.

[114] S. Wakabayashi, *Remarks on hyperbolic polynomials*, Tsukuba J. Math. **10** (1986), no. 1, 17–28.

[115] H. Weyl, *Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung)*, Math. Ann. **71** (1912), 441–479 (German).

[116] Y. Yomdin, *The zero set of an "almost polynomial" function*, Proc. Am. Math. Soc. **90** (1984), no. 4, 538–542.

Index

Abhyankar–Jung theorem, 11
absolutely continuous, 56, 59
AC, 59
$\mathcal{A}_d(\mathbb{R}^m)$, 36

basic invariants, 40, 47
  minimal system of, 40, 47
bounded variation, 59
Bronshtein's theorem, 15, 18
BV, 59

Cartan subspace, 42, 49
categorical quotient, 41, 47
characteristic map, 28
characteristic roots, 28
χ_A, 6
C^∞, 57
C^k, 57
$C^{k,\alpha}$, 58
C^ω, 57
coregular representation, 48
$\mathbb{C}[V]^G$, 47

DC, 28
definable, 26
Δ_A, 6
Δ_P, 6
$\mathrm{Diag}(d)$, 6, 45
difference-convex, 28

elementary symmetric functions, 40

$\mathcal{F}^\beta$, 25
fixed points, 43
f_{red}, 6

geometric quotient, 47
Glaeser's inequality, 20
$\mathrm{GL}_n(\mathbb{C})$, 6
$G \circlearrowright V$, 40, 47
$(G \circlearrowright V, d, \sigma)$, 41, 47
Gårding hyperbolic polynomial, 26

$\mathcal{H}$, 57
Hölder–Lipschitz space, 58
$\mathrm{Herm}(d)$, 6, 29
higher order Glaeser inequalities, 35
$\mathcal{H}^k$, 5
$\mathrm{Höld}_{\alpha,\Omega}(f)$, 58
$\mathrm{Hyp}(d)$, 6, 60
hyperbolic polynomial, 2
$\mathrm{Hyp}^0_T(d)$, 6, 9
$\mathrm{Hyp}_T(d)$, 6, 9, 61

inner Wasserstein distance, 56
isotropy subgroup, 44

Lax conjecture, 27
Lebesgue space, 58
lifting, 40, 41, 47
linearly reductive group, 47
$\mathrm{Lip}_\Omega(f)$, 58
$\mathcal{L}^n$, 5
local blowings-up, 43
local power substitution, 51
L^p, 58
L^p_w, 35, 58

$\mathrm{Mat}_{m,n}(R)$, 6
$\mathrm{Mat}_n(R)$, 6
metric speed, 56
multiplicity, 25
multivalued Sobolev functions, 36

Newton identities, 8
Newton polynomials, 8
normally nonflat, 25, 43

o-minimal, 26
$\mathrm{O}_n(\mathbb{R})$, 6
orbit map, 41
orbit space, 41, 47
orbit type, 45
  principal, 45
orbit type stratification, 45

parameterization of multivalued function, 36
parameterization of the roots, 6
polar representation, 42, 49
polynomial
  elementary symmetric, 40
  Gårding hyperbolic, 26
  hyperbolic, 2
  Newton, 8
  quasi-ordinary, 11
  real stable, 27
  stable, 27
pseudoreflection, 48

quasi-ordinary polynomial, 11

real stable polynomial, 27
Rellich's theorem, 7
$\mathbb{R}[V]^G$, 40

SBV, 60
section, 42, 49
sign of a function, 54
singular value decomposition, 13
singular values, 13, 31
slice representation, 44
slice theorem, 44
S_n, 6
Sobolev space, 59
special functions of bounded variation, 60

splitting, 9, 19
stable point, 49
Sym(d), 6, 45
system of the roots, 6

theorem
  Abhyankar–Jung, 11
  Bronshtein's, 15, 18
  Hermite–Sylvester version of Sturm's, 61
  Rellich's, 7
  slice, 44
Tschirnhausen form, 8
Tschirnhausen transformation, 8

$\mathrm{U}_n(\mathbb{C})$, 6
unordered tuple, 6, 36

Vandermonde manifold, 60
Vandermonde mapping, 60
$V /\!\!/ G$, 41, 47
V/G, 41, 47
V^G, 43

Wasserstein distance, 55
  inner, 56
weak Lebesgue space, 35, 58
Weyl group, 42, 49
$W^{k,p}$, 59

Adam Parusiński: Université Côte d'Azur, CNRS, LJAD, UMR 7351, 06108 Nice, France
Email address: adam.parusinski@univ-cotedazur.fr

Armin Rainer: Fakultät für Mathematik, Universität Wien, Oskar-Morgenstern-Platz 1, A-1090 Wien, Austria
Email address: armin.rainer@univie.ac.at