A REAL ANALYTIC APPROACH TO ESTIMATING OSCILLATORY INTEGRALS

Maxim Gilula

A DISSERTATION

in

Mathematics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

____

Philip T. Gressman, Professor of Mathematics

Graduate Group Chairperson

____

David Harbater, Professor of Mathematics

Dissertation Committee:
Philip T. Gressman, Professor of Mathematics
Robert M. Strain, Professor of Mathematics
Tony Pantev, Professor of Mathematics

Dedicated to my grandfather, Max Davidovich Gilula, and to my parents, Mikhail and Natalia.

Acknowledgments

First I would like to thank my parents Mikhail and Natalia, my sister Victoria, and my grandfather Max, for caring about my education and helping me when fractions, and later logarithms, didn't make any sense. Without them I surely wouldn't be where I am today.

I am lucky to have been a student in Mr. Perry Carter's exciting algebra class, and grateful to him for his suggestion to get ahead in math over summer. I certainly learned more from Mrs. Sheila Scafine and her calculus courses than I could from a college course with the same syllabus, and would also like to thank her for encouraging participation in math competitions, which helped me see that math was even more fun than I already thought. Thank you both for setting me on the right path to becoming a math major.

In my first year of undergrad, I read *Journey Through Genius* by Dr. William Dunham, from which my inspiration for math grew tenfold; thank you for writing this wonderful book. I appreciate Dr. Sarah Eichhorn immensely for opening my eyes to what undergraduate research is about, and for all of her encouragement. I am very grateful to Dr. Martin Zeman for all the time and effort he spent with me while I was trying to study set theory, and for helping me get through reading my first academic papers. Thank you so much, Damek Davis, for showing me what classes a student who plans to go to graduate school should be taking, and for being a great inspiration. I want to thank you all for setting me on the right path to becoming a graduate student.

I would like to deeply thank my adviser, Professor Philip T. Gressman, for being a great teacher and adviser. He has always been kind, patient, and welcoming. I am very grateful that he suggested a research problem that was very fun to work on for me (found below), and that has so much potential to being expanded in my future research. I am very grateful to the rest of the analysis group for all of the enlightening student seminars we organized and learned from, in particular to Soumyashant Nayak for studying harmonic analysis with me when there was no course available at that time, and more. I am also grateful to Martin Citoler-Saumell and others who were always there to listen to my ideas and provide suggestions. Thank you all for setting me on the right path to becoming a mathematician.

I am grateful to Professor Michael Greenblatt for helping me with all the questions I had about his work, as well as reading an early version of this thesis and discussing with me some of the ideas contained within.

I am also grateful to Professor Joe Kamimoto for discussing his work with me, as well as helping me with an explicit computation of some exponents in the asymptotic expansion found in his recent paper with Professor Toshihiro Nose.

I learned a vast amount from Temple's graduate student analysis seminar; thank you Eric Stachura for inviting me to be a part of your group. Also, thank you for the enormous amount of editing you helped me with.

Thank you Alanna Smothers for proofreading many of my essays and everything else you have done for me; I had many of my best ideas during holidays spent at your home. The love and support you provided made a huge difference in my life the past few years.

Finally I would like to thank Larry Lederer, who has been my tennis partner consistently for the past few years. He has been helping me continue my other passion – playing tennis at a high level – and I can't thank him enough for that. Thank you also for just being an overall champion of encouragement.

Thank you so much, all of my great family, friends, teachers, and Professors, and so many others to whom I owe my gratitude.

ABSTRACT

A REAL ANALYTIC APPROACH TO ESTIMATING OSCILLATORY INTEGRALS

Maxim Gilula

Philip T. Gressman

We develop an asymptotic expansion for oscillatory integrals with real analytic phases. We assume the phases satisfy a nondegeneracy condition originally considered by Varchenko, which is related to the Newton polyhedron. Analogous estimates for smooth and C^k phases are also proved. With algebraic techniques such as resolution of singularities, Varchenko was the first to obtain sharp estimates for oscillatory integrals with nondegenerate analytic phases, assuming the Newton distance of the phase is greater than 1. This condition has also been frequently used in modern literature; for example, Greenblatt and later Kamimoto-Nose obtained more general results by also using resolution of singularities. Using only real analytic methods that are very much in the spirit of van der Corput, we develop a full asymptotic expansion for analytic phases satisfying Varchenko's condition, and an asymptotic expansion with finitely many terms for C^k phases under the additional assumption that the Newton polyhedron intersects each coordinate axis. We demonstrate how the exponents in the asymptotic expansion of these integrals can be obtained completely geometrically via the Newton polyhedron. Important techniques include: dyadic decomposition;

proving and then using a lower bound similar to that of Lojaciewicz for analytic functions, together with the method of stationary phase to integrate by parts; linear programming to get sharpest estimates (matching Varchenko's); and finally, repeated differentiation of the integral with respect to the oscillatory parameter in order to obtain higher order terms of the expansion.

Contents

Chapter 1

Introduction

The main goal of this thesis is to use real analytic methods to develop an asymptotic expansion for scalar oscillatory integrals with analytic, smooth, and C^k phases. Namely, integrals of the form

$$I(\lambda) = \int_{\mathbb{R}^d} e^{i\lambda\phi(x)}\psi(x)dx \tag{1.0.1}$$

are studied, where $\phi : \mathbb{R}^d \to \mathbb{R}$ is called the phase, and the amplitude ψ is supported close to the origin and enjoys the same smoothness as ϕ. The nondegeneracy conditions we assume on the phase vary depending on its smoothness. It has been known since at least the 1970's that, as $\lambda \to \infty$, the integral $I(\lambda)$ admits an asymptotic expansion of the form

$$I(\lambda) \sim \sum_{p} \sum_{r=0}^{d-1} a_{p,r}(\psi)\lambda^{-p} \log^{d-1-r}(\lambda), \tag{1.0.2}$$

where p runs through finitely many arithmetic progressions, independent of ψ, constructed from positive rational numbers. One can find this, for example, in Malgrange [15]; a more modern proof can be found in Greenblatt[5]. To deal with the integrals under consideration, many authors use the Newton polyhedron of the phase as an important tool. The Newton polyhedron contains information about the growth of polynomials. If the Newton polyhedron of the phase or the amplitude intersects each coordinate axis, Kamimoto-Nose found a set containing the exponents p for C^∞ functions, as well as for analytic functions with no additional condition on the Newton polyhedron (see [9, Theorem 11.1]). The methods of Kamimoto-Nose were mainly algebraic and complex analytic, as they made use of techniques related to toric resolution and computed poles of integrals reminiscent of [15]. Most proofs in the literature making use of the Newton polyhedron to prove sharp estimates for $I(\lambda)$ involve the use of resolution of singularities, e.g., [9], [4], [18]. Unfortunately, resolution of singularities is very difficult in high dimensions, and is problematic for analysts who wish to not change the coordinates. Vasiliev avoided resolution of singularities by instead using techniques from complex algebraic geometry in [19] to obtain asymptotic behavior for C^∞ phases having an absolute minimum. Our proof is completely real analytic and with it we are able to develop an expansion for nondegenerate C^k

functions assuming the Newton polyhedron intersects each coordinate axis, as well as an analogous result for real analytic functions only satisfying nondegeneracy.[1] The use of real analytic methods allows us, in particular, to have **no change of the original coordinate system** for any of our results. Real analytic methods allow for more analysts to use Varchenko's result without being forced to use it as a "black box" because of the algebraic techniques used. The best feature of the proof may be that it is very close in spirit to modern proofs of van der Corput's lemma. We make use of the method of stationary phase to obtain estimates for integrals away from singularities, then optimize the exponent of λ just like in modern proofs of van der Corput's lemma for C^k functions of one variable. However, we are unable to prove the sharpness of the estimate without considering explicit examples of phases (also analogous to van der Corput). For this reason, we simply cite the sharpness by Varchenko's result. The exponents p as well as the exponents $d-1-r$ of the log term in (1.0.2) we obtain are easy to visualize geometrically and we provide examples after the statement of Theorem 2. We believe that the methods in this thesis are more robust than previous methods, and are particularly well-suited for a very important yet difficult question in this subject that has very little progress in the literature: the stability of estimates of $I(\lambda)$ under perturbations of the phase. More will be discussed about this in the last chapter.

The next chapter begins by introducing some notation that is used throughout the

[1] The method of proof also works for phases that can be expressed as a Taylor series with remainder, replacing all exponents $\alpha \in \mathbb{N}^d$ with exponents $\alpha \in \mathbb{N}^d/k = \{n/m : n \in \mathbb{N}\}$.

document. Some of it is not conventional, but is very useful for the computations we want to perform, e.g., for $x, y \in \mathbb{R}^d$ we define $xy = (x_1y_1, \ldots, x_dy_d)$. This notation, along with a parametrization of $(0, 1)$ by supporting hyperplanes introduced later on, will help us understand exactly why the Newton polyhedron should be considered for x close enough to the origin. Once most of the notation and preliminary discussions are complete, we move on to the main results. The main results consist of two lemmas and two theorems, with Theorem 2 as the main attraction.

Chapter 3 introduces the Newton polyhedron $N(\phi)$ for C^k functions $\phi : \mathbb{R}^d \to \mathbb{R}$ and introduces a parametrization of supporting hyperplanes that is used throughout the text. If ϕ is not real analytic, we need to assume that if $\phi = P_m + R_m$, where P_m is a polynomial of order m and R_m is the remainder, then $N(P_m)$ intersects each coordinate axis. The assumption that the Newton polyhedron intersects each coordinate axis is a very natural condition that can be found in many papers about oscillatory integrals, e.g., [9], [7], [19], and many more. [2] The statement of Theorem 2 from the previous chapter may take a few moments to absorb, so we illustrate the simplicity of what the theorem is saying with a few examples at the end of Chapter 3. We compute the first few terms of the asymptotic expansion for a C^k function and in another example we will consider a real analytic function with asymptotic expansion showing the main aspects of how the geometry of the Newton polyhedron affects the asymptotic behavior of $I(\lambda)$. The description of the exponents in the asymptotic

[2] In particular, a generic polynomial of any degree satisfies this condition.

expansion will have to wait until some more notation is introduced. Fortunately, the exponents can be described geometrically.

The results are proved below in the order they are stated, starting with the proof of Lemma 1 in Chapter 4. This chapter begins with an intuitive illustration of how supporting hyperplanes of the Newton polyhedron govern the behavior of the phases under consideration. The body of this chapter consists of proving preliminary linear algebraic results, culminating with the main proposition– Proposition 6 says that if we have a supporting hyperplane H that is "too close"to a face $F \subset N(\phi)$ not contained in H, then we can move to a new hyperplane containing F at the cost of rescaling the domain we are estimating over by a uniform constant. In this case, F is necessarily lower codimension than the face we started with, so after at most $d-1$ steps this process must terminate and provide a hyperplane that is good in the sense we require. This can be viewed as some type of compactness theorem. Once this proposition is proved, we are able to prove Lemma 1. In this chapter (Equation 4.3.12) we define what it means for $\varepsilon \in (0,1)^d$ to be small enough; this precise statement requires some of the preliminary results contained in Chapter 4, so we do not provide it in the statement of Lemma 1 in Chapter 2. This result is closely related to Lojaciewicz's famous lower bound for analytic functions evaluated away from their zero set[14]. Very similar results have also been considered by Fukui-Yoshinaga[2] and Yoshinaga[20]. Both of these results use resolution of singularities. A connection between Lemma 1 and the exponent in Lojaciewicz's lower bound can be found in Abderrahmane[1], assuming the origin is only an isolated singularity of the phase. This chapter is very important because it is here that we bypass using resolution of singularities for nondegenerate functions with only linear algebra.

In Chapter 5, a dyadic partition of unity is used to split $I(\lambda)$ into a sum of dyadic pieces that are each easy to estimate; here we only obtain estimates of $I(\lambda)$ over the first orthant without loss of generality; the argument is identical over each orthant. We note that because of the dyadic decomposition, the cancellation between orthants potentially making the estimate of $I(\lambda)$ sharper goes unnoticed and could be one of the factors for why we do not generally obtain sharp estimates for $t \leq 1$. We do not know whether our estimate is sharp over each orthant. However, the tools applied are powerful enough to provide estimates as sharp as Varchenko's for all $t > 0$[18]. In particular, they are sharp for $t > 1$. The main tools used in the estimates are integration by parts and Lemma 1.

Next, we use linear programming to prove Theorem 1. Varchenko's estimates easily fall out as a special case. We first prove a simple case in order to highlight the methods being used and how exactly the geometry of $N(\phi)$ affects the estimate guaranteed by Theorem 1.

Finally, we prove Theorem 2. The proof goes by obtaining an estimate for $(\lambda\frac{d}{d\lambda} + 1/t)I(\lambda)$, then applying more differential operators to $I(\lambda)$ and obtaining estimates by induction. We then obtain a differential inequality for which we require a simple ODE result. This allows us to find the first N terms of $I(\lambda)$, where N depends on the smoothness of ϕ and ψ.

We conclude with how the results in this document could be applied to future work. in the appendix, we work out a couple of examples from [18] and in particular check that Theorem 2 predicts the exponents appearing in the asymptotic expansion of $I(\lambda)$ when the phase has Newton distance less than 1.

Chapter 2

Main results

2.1 Conventions and terminology

2.1.1 Basic notation

Throughout this text, we use the following conventions:

- $\mathbb{N}$ is the set of nonnegative integers,
- $\mathbb{R}_{\geq}$ is the set of nonnegative reals,
- $f'_{x_i}(x) = \frac{\partial}{\partial x_i} f(x)$,
- the inner product of $x, y \in \mathbb{R}^d$ is $\langle x, y \rangle = x_1 y_1 + \cdots + x_d y_d$; we also use $x \cdot y$ for the inner product, usually when functions are involved,
- the notation $\cdot^\gamma$ is used for the operator defined by $(\cdot^\gamma f)(x) = x^\gamma f(x)$,

- $\|v\|_m$ of of a vector $v \in \mathbb{R}^d$ (or a matrix $v \in \mathbb{R}^{d\times d}$) is the standard ℓ^m norm of the vector v for $1 \leq m \leq \infty$,

- and for $0 \leq k < \infty$, C^k represents the class of k times continuously differentiable functions. The domain of the functions is clear through context. Moreover, we refer to C^∞ functions as smooth.

If x is a $d-$tuple we write $x = (x_1, \ldots, x_d)$. In particular, subscripts denote components of a vector. On the other hand, whenever we have a list of $d-$tuples, they are indexed by a superscript, e.g., $\{\alpha^i\}_{i\in I}$. There is one consistent exception: the standard unit normals $\mathbf{e}_i \in \mathbb{R}^d$ defined componentwise by $\mathbf{e}_{ij} = \delta_{ij}$ (the Kronecker delta).

Next, some algebraic conventions are introduced. In addition to the standard notation, for $y \in \mathbb{R}^d_\geq$ and $\alpha \in \mathbb{R}$, that $\partial^\alpha = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_d}}{\partial x_d^{\alpha_d}}$, the exponentiation of vectors $y^\alpha = y_1^{\alpha_1} \cdots y_d^{\alpha_d}$, as well as $|y| = y_1 + \cdots + y_d$, we make use of some less standard notation for $c \in \mathbb{R}$ and $y, z \in \mathbb{R}^d_\geq$:

- $yz = (y_1 z_1, \ldots, y_d z_d)$,

- if $c > 0$, denote the vector $(c^{y_1}, \ldots, c^{y_d})$ by c^y,

- if $c > 1$, $[y, cy]$ is defined to be the box $\prod_{j=1}^d [y_j, cy_j]$,

- if the components of y are positive, $y^{-z} = y_1^{-z_1} \cdots y_d^{-z_d}$,

- if the components of y are positive, $f_y(x) = f(y^{-1}x)$, and

- boldface $\mathbf{c}$ denotes the vector $(c, \ldots, c)$.

In particular, note that $(c^y)^z = c^{\langle y,z\rangle}$ and $(c^y x)^z = c^{\langle y,z\rangle} x^z$. Also note that $x\nabla\phi(x) = (x_1\phi'_{x_1}(x), \ldots, x_d\phi'_{x_d}(x))$. Since we do not have a notion of a vector raised to a constant, there is no ambiguity in indexing vectors by superscripts. Lastly, we write

$$f(x) \lesssim g(x)$$

for positive real-valued functions f and g to express that there is a positive constant C independent of x such that $f(x) \leq Cg(x)$ for all x wherever this expression makes sense. There may be multiple variables in the domain of f and g, and we state explicitly the independence of the implicit constant whenever we use this notation.

2.1.2 The context for the rest of the document

Although the main goal is to estimate $I(\lambda)$ in (1.0.1) with domain of integration all of $\mathbb{R}^d$, in this thesis we mainly consider estimates over the first orthant. There is no loss of generality because given $I(\lambda)$, we can split the integral into 2^d integrals, one for each orthant, and approximate each separately. Because of this symmetry, we assume the amplitude ψ is supported in a neighborhood of the origin, but take the integral only in the orthant containing those $d-$tuples with all nonnegative components. To summarize, the goal up to and including Theorem 1 is to provide estimates for the integral $I_+(\lambda)$ defined by

$$I_+(\lambda) = \int_{\mathbb{R}^d_\geq} e^{i\lambda\phi(x)}\psi(x)dx, \tag{2.1.1}$$

where ϕ is real analytic in $[-4,4]^d$ and ψ is smooth and supported close enough to the origin in the set $[-4,4]^d$, or else both are $C^k([-4,4]^d)$ with the extra assumption that the Newton polyhedron of ϕ intersects each coordinate axis (see Chapter 3). The number 4 above is for convenience: it can be replaced by any positive real number we wish. **Without loss of generality we assume that $\phi(\mathbf{0}) = 0$, but is not identically zero in any neighborhood of the origin, and $\nabla\phi(\mathbf{0}) = 0$ for the rest of the document.** If $\phi(\mathbf{0}) = 0$, we could factor out $e^{i\lambda\phi(\mathbf{0})}$ and consider the phase $\phi - \phi(\mathbf{0})$, and if $\nabla\phi(\mathbf{0}) \neq 0$, the estimates obtained are trivial by the method of stationary phase.

To estimate $I_+(\lambda)$, we use a partition of unity and reduce the problem to estimating

$$I_+(\lambda, \varepsilon) = \int_{[\varepsilon, 4\varepsilon]} e^{i\lambda\phi(x)}\eta_\varepsilon(x)\psi(x)dx$$

where $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_d) \in (0,1)^d$ is small enough (see 4.3.12), $[\varepsilon, 4\varepsilon]$ is the box $\prod_{j=1}^d [\varepsilon_j, 4\varepsilon_j]$, and η is smooth with support in $[1,4]^d$. Here, $\{[\varepsilon, 4\varepsilon]\}$ is a set of dyadic boxes and $\{\eta_\varepsilon\}$ is a smooth partition of unity of $(0,1)^d$. We decompose our amplitude

ψ into a sum of amplitudes $\eta_\varepsilon\psi$ supported in $[\varepsilon, 4\varepsilon]$, estimate each $I_+(\lambda, \varepsilon)$, and sum these estimates to estimate $I_+(\lambda)$. In order to discuss the main results, we need to introduce the Newton polyhedron – it contains the information necessary to determine which monomials of a given polynomial are largest near the origin. Together with Varchenko's nondegeneracy condition[18, Definition 5], the Newton polyhedron gives us information about the largest terms in the Taylor expansion of ϕ.

2.2 Statements of the Results

Although we will define everything rigorously in the following chapter, we quickly present some definitions necessary for stating the main results.

For an analytic function $\phi(x) = \sum_\alpha c_\alpha x^\alpha$ defined in a neighborhood of the origin, we define the Newton polyhedron of ϕ, denoted $N(\phi)$, to be the convex hull of

$$\bigcup_{c_\alpha \neq 0} \alpha + \mathbb{R}^d_\geq .$$

Also, we say that ϕ is nondegenerate if for all x satisfying $x_1 \cdots x_d \neq 0$, there is some $1 \leq i \leq d$ such that $x_i \phi'_{x_i}(x) \neq 0$.

If ϕ is C^k in a neighborhood of the origin, we define $N(\phi) = N(P_k)$, where P_k is the Taylor polynomial of order at most k. Finally, we say ϕ is convenient if it is C^k in a neighborhood of the origin and $N(P_k)$ intersects each coordinate axis. More explicitly, for any $m \leq k$ we say ϕ is $m-$convenient if P_m intersects each coordinate axis. If ϕ is convenient, we say ϕ is nondegenerate if P_k is nondegenerate.

A crucial step in proving the main theorems is quantifying how $\nabla\phi$ behaves near the origin:

Lemma 1. *Assume ϕ is analytic or convenient in a neighborhood of the origin. Assume ϕ is nondegenerate. For all $\varepsilon \in (0,1)^d$ small enough, for all x in the box $[\varepsilon, 4\varepsilon]$, and for all $\alpha \in N(\phi)$, we have the lower bound*

$$\|x\nabla\phi(x)\| \gtrsim \varepsilon^{\alpha},$$

where the implicit constant is independent of ε.

So we see nondegeneracy implies the sharpest possible growth rate for $\nabla\phi$ around the origin. Small enough is made explicit in (4.3.12). With this lemma we are able to prove the most useful result in the thesis:

Lemma 2. *Let $\beta \in \mathbb{N}^d$. Let ϕ be C^k nondegenerate and convenient in a neighborhood of the origin. Assume $\eta : \mathbb{R}^d \to \mathbb{R}$ is C^k with support in $[1,4]^d$. For all $\varepsilon \in (0,1)^d$ small enough, we have the estimate*

$$\Big|\int_{\mathbb{R}^d} e^{i\lambda\phi(x)} x^{\beta} \eta_{\varepsilon}(x)dx\Big| \lesssim \lambda^{-N}\varepsilon^{-(N\alpha-\beta-\mathbf{1})} \tag{2.2.1}$$

for all $\lambda > 0$, all $0 \leq N \leq k$, and all $\alpha \in N(\phi)$, where the implicit constant above is independent of ε and λ.

If in addition ϕ and η are smooth, then the estimate (2.2.1) holds for all $0 \leq N < \infty$.

If instead ϕ is real analytic (not necessarily convenient) and η is smooth, the estimate (2.2.1) holds for all $0 \leq N < \infty$.

With the help of Lemma 2, we prove a useful generalization of Varchenko's upper bounds. Below we use the notation $w(\beta + \mathbf{1})$, which is explained in the following chapter. However, we can say that $d_j + 1$ below (the next two theorems) is the greatest codimension over any face containing $\beta + \mathbf{1}$. Also, $c = \langle \beta + \mathbf{1}, w(\beta + \mathbf{1}) \rangle$ is such that $(\beta + \mathbf{1})/c$ is contained in the boundary of $N(\phi)$.

Theorem 1. *Let ϕ be C^k nondegenerate and convenient in a neighborhood of the origin. Let $\beta \in \mathbb{N}^d$ be such that $\langle \beta + \mathbf{1}, w(\beta + \mathbf{1}) \rangle < k$. If $\psi : \mathbb{R}^d \to \mathbb{R}$ is C^k and supported close enough to the origin, there is a uniform constant independent of λ such that*

$$\left| \int_{\mathbb{R}^d} e^{i\lambda\phi(x)} x^\beta \psi(x) dx \right| \lesssim \lambda^{-\langle \beta + \mathbf{1}, w(\beta + \mathbf{1}) \rangle} \log^{d_j}(\lambda) \tag{2.2.2}$$

where $d_j = \min\{d, |w(\beta + \mathbf{1})|\} - 1$.

If in addition ϕ and ψ are smooth, then (2.2.2) holds for all $\beta \in \mathbb{N}^d$.

If instead ϕ is real analytic (not necessarily convenient) and ψ is smooth, (2.2.2) holds for all $\beta \in \mathbb{N}^d$.

For the last theorem, we claim there is a well-ordered set $\mathcal{C}$ containing the exponents for the asymptotic expansion of $I(\lambda)$. This is discussed at the end of Chapter 3.

Theorem 2. *Assume $\phi : \mathbb{R}^d \to \mathbb{R}$ is C^k nondegenerate in a neighborhood of the origin. Assuming $N(\phi)$ exists, let $p_0 < p_1 < \cdots$ be the ordering of $\mathcal{C}$. Let $K_j = \max\{|w(\alpha)| : \langle \alpha, w(\alpha)\rangle = p_j\}$ and define $0 \le d_j \le d-1$ by $d_j = \min\{d, K_j\} - 1$. Order the set*

$$\{\lambda^{-p_j} \log^r(\lambda) : j \in \mathbb{N}, 0 \le r \le d_j\} = \{E_\ell(\lambda) : \ell \in \mathbb{N}\}$$

so that for all $n \in \mathbb{N}$ we have $E_{n+1}(\lambda)/E_n(\lambda) \to 0$ as $\lambda \to \infty$.[3]

(i) Assume in a neighborhood of the origin ϕ is either real analytic or convenient smooth. Let $\psi : \mathbb{R}^d \to \mathbb{R}$ be smooth and supported close enough to the origin. Then, there are constants $a_\ell(\psi) \in \mathbb{C}$ such that for all $N \in \mathbb{N}$,

$$\left| \int_{\mathbb{R}^d_{\geq}} e^{i\lambda\phi(x)}\psi(x)dx - \sum_{\ell=0}^{N} a_\ell E_\ell(\lambda) \right| \lesssim E_{N+1}(\lambda). \tag{2.2.3}$$

for all λ large enough. The implicit constant is independent of λ.

(ii) Assume ϕ is $m-$convenient and $\psi : \mathbb{R}^d \to \mathbb{R}$ is C^k with support close enough to the origin. Assume $k > d(2m+d)(n+1)+d$.[4] *Then there are constants*

[3] To define everything so far, all we needed was a polyhedron. In particular, we did not need any smoothness assumptions on ϕ up to this point.

[4] This is nowhere near a sharp lower bound on k. A sharper bound will be discussed in future work.

$a_{j,r}(\psi) \in \mathbb{C}$ such that

$$\int_{\mathbb{R}^d} e^{i\lambda\phi(x)}\psi(x)dx \sim \sum_{j=0}^{r}\sum_{r=0}^{d_j} a_{j,r}(\psi)\lambda^{-p_j}\log^{d_j-r}(\lambda)$$

in the sense of 2.2.3. Moreover,

$$\left|I(\lambda) - \sum_{j=0}^{n}\sum_{r=0}^{d_j} a_{j,r}(\psi)\lambda^{-p_j}\log^{d_j-r}(\lambda)\right| \lesssim \lambda^{-p_{n+1}}\log^{d-1}(\lambda)$$

for all λ large enough. The implicit constant is independent of λ.

Chapter 3

The Newton polyhedron

3.1 What the Newton polyhedron represents

In *Lectures on polytopes*, Ziegler[21] defines an $\mathcal{H}-$**polyhedron** as an intersection of finitely many halfspaces in some $\mathbb{R}^d$. We refer to this simply as a polyhedron. A **face** F of a polyhedron P is a subset of P that can be written as $F = H \cap P$ for some hyperplane $H = \{\xi \in \mathbb{R}^d : \langle \xi, w \rangle = b\}$, where $w \in \mathbb{R}^d$ and $b \in \mathbb{R}$ are fixed. We say $v^0, \ldots, v^k$ are **affinely independent** if $v^1 - v^0, \ldots, v^k - v^0$ are linearly independent for $k \neq 0$. We say more precisely that a face F is a **dimension** $\boldsymbol{k}$ face in $\mathbb{R}^d$ (or **codimension** $\boldsymbol{d-k}$) when we can find a set of $k+1$ but not $k+2$ vectors in F that are affinely independent if $k > 0$, and say the face F is **dimension** 0, or a **vertex**, if $F = \{v\}$ for some $v \in P$.

As stated in the previous chapter, the Newton polyhedron of a given polynomial contains all of the necessary information in order to determine which monomials could be the largest.[5] Given a supporting hyperplane in two dimensions, for example, we can parametrize it by $u/a + v/b = 1$ for $a, b \in (0, \infty)$ if the hyperplane does not contain the origin. In this case the hyperplane represents when $x^a = y^b$. In particular, we can determine how to scale $x^a + y^b$ homogeneously. We can conclude that such a scaling forces x^a and y^b to be the largest for x, y near the origin. More generally, the extreme points of the Newton polyhedron represent the powers of the monomials that can be largest, and faces (of any dimension) show which monomials can possibly be comparable. So the Newton polyhedron should be thought of as a generalization of degree in the sense of limiting behavior. For example, in one dimension we know exactly when $x^k < x^m$ for positive x near the origin: when $m < k$. But in higher dimensions, degree has little to do with which term is largest: the monomial x^{53} can be greater than xy for positive small x, y (e.g., consider the region $x^{52} > y$.) Although the Newton polyhedron contains information about the largest monomials, cancellation can occur and throw off the intuition: if we let $\phi(x, y) = x - y + x^2$, indeed either x or y are the largest away from $x = y$. If $x = y$, then x^2 is the largest. Roughly speaking, Varchenko's nondegeneracy condition guarantees that no cancellation of this form can happen in each component of $\nabla\phi$.

Not only is the Newton polyhedron the most important object of study in the proof of Lemma 1, which is natural because this lemma tells us how a nondegenerate

[5]We define the Newton polyhedron in order to obtain information about small values of x. One could define it so that it provides information about large values, or information about only small x_1, etc.

function behaves near the origin, it provides us an easy way to geometrically visualize the exponents appearing in the main result of this thesis: Theorem 2, the asymptotic expansion of $I(\lambda)$.

We first define the Newton polyhedron for analytic functions defined in a neighborhood of the origin. Then we move on to m times continuously differentiable functions with a condition that guarantees the Newton polyhedron captures the information we are interested in.

3.2 The Newton polyhedron and nondegeneracy

As a reminder, we always assume ϕ is defined in a neighborhood of the origin satisfying $\phi(\mathbf{0}) = 0$ and ϕ not identically zero. We begin with a definition.

Definition 1 (Taylor support). *Let $\phi : \mathbb{R}^d \to \mathbb{R}$ be analytic in a neighborhood of the origin. Denote the set of indices of the nonzero coefficients in the unique expansion $\phi(x) = \sum_\alpha c_\alpha x^\alpha$ by*

$$\operatorname{supp}(\phi) = \{\alpha \in \mathbb{N}^d : c_\alpha \neq 0\}$$

and call $\operatorname{supp}(\phi)$ *the* ***Taylor support*** *of* ϕ.[6]

The set $\operatorname{supp}(\phi)$ aids us in defining the Newton polyhedron of ϕ :

Definition 2 (Newton polyhedron of an analytic function). *Let $\phi : \mathbb{R}^d \to \mathbb{R}$ be*

[6]Whenever we refer to the set of inputs x of a function f where $f(x)$ is nonzero, we simply use the word "support,"and to avoid ambiguity, whenever we talk about the Taylor support of f, we write $\operatorname{supp}(f)$.

analytic defined in a neighborhood of the origin. We define the ***Newton polyhedron*** *of ϕ to be the convex hull of the union*

$$\bigcup_{\alpha \in \operatorname{supp}(\phi)} \alpha + \mathbb{R}^d_{\geq},$$

and we denote the Newton polyhedron of ϕ by $N(\phi)$.

Now that we have defined the Newton polyhedron, we show that it is indeed a polyhedron, and make heavy use of the fact that it has finitely many codimension 1 faces throughout all of the main results. To show $N(\phi)$ is a polyhedron whenever ϕ is analytic, we take the route of showing that $N(\phi)$ has finitely many extreme points.

Proposition 1. *Let $\phi(x) = \sum c_\alpha x^\alpha$ be real analytic in a neighborhood of the origin in $\mathbb{R}^d$. Then $N(\phi)$ has finitely many extreme points.*

Proof. First, we show that if $N(\phi)$ has any extreme points, they must lie in $\operatorname{supp}(\phi)$. Let β be an extreme point of $N(\phi)$. Since $N(\phi)$ is the convex hull of the union of sets of the form $\alpha + \mathbb{R}^d_{\geq}$ where $\alpha \in \operatorname{supp}(\phi)$, we see all $\beta \in N(\phi)$ can be expressed as some convex combination

$$\beta = \sum_{i=1}^{n} \lambda_i(\alpha^i + \gamma^i),$$

where $\lambda_i' s$ are positive reals summing to 1, $\alpha^i \in \operatorname{supp}(\phi)$ are distinct, and $\gamma^i \in \mathbb{R}^d_{\geq}$. If β is an extreme point, then $\beta = \alpha + \gamma$ since it cannot be a convex combination of distinct points. Next, since we can express $\alpha + \gamma$ as the convex combination $\alpha/2 + (\alpha + 2\gamma)/2$

of vectors α and $\alpha + 2\gamma$ lying in $N(\phi)$, we see that necessarily $\gamma = \mathbf{0}$, and conclude $\beta = \alpha + \gamma = \alpha$.

Next, we show that there must be finitely many extreme points. If there are infinitely many, we can consider an enumeration $\{\alpha^i\}_{i\in\mathbb{N}}$ of all extreme points α of $N(\phi)$. There are indeed at most countably many because $\mathrm{supp}(\phi) \subseteq \mathbb{N}^d$ is at most a countable set.

Observe that for each $j \neq k$ there is some $1 \leq \ell \leq d$ such that $\alpha^k_\ell < \alpha^j_\ell$ or else we would have $\alpha^k = \alpha^j + \gamma$ for some $\gamma \in \mathbb{R}^d_{\geq}$. Therefore, there is some $1 \leq \ell \leq d$ and a subsequence $\{\alpha^{n_i}\} \subset \{\alpha^i\}$ of extreme points satisfying the infinite chain of strict inequalities

$$\alpha^{n_1}_\ell > \cdots > \alpha^{n_j}_\ell > \cdots .$$

Since $\mathbb{N}$ is well-ordered, this cannot be the case. Therefore there is no such subsequence and we must have only finitely many extreme points. $\square$

Theorem 1.2 in Ziegler[21] states that any polyhedron is a Minkowski sum of a convex hull of a finite set of points plus a conical combination of vectors. In this case, we want to consider the cones $\mathbb{R}^d_{\geq}$. Recall that the **Minkowski sum** $P + Q$ of two sets $P, Q \subset \mathbb{R}^d$ is defined by

$$P + Q = \{p + q : p \in P \text{ and } q \in Q\}.$$

In our case, we know that $N(\phi)$ is the sum of the convex hull of the finitely many extreme points plus the cone $\mathbb{R}^d_{\geq}$. This is because, using the notation of Proposition 1, by definition of the Newton polyhedron we know any vector β in $N(\phi)$ can be written as a convex combination

$$\beta = \sum_{i=1}^{n} \lambda_i(\alpha^i + \gamma^i) = \sum_{i=1}^{n} \lambda_i \alpha^i + \sum_{i=1}^{n} \lambda_i \gamma^i.$$

The right hand side shows that β is in the Minkowski sum we are considering.

In order to define nondegeneracy of ϕ, we need to consider the following polynomials.

Definition 3 (The polynomials ϕ_F)**.** *Let $\phi : \mathbb{R}^d \to \mathbb{R}$ be real analytic. For any compact face $F \subset N(\phi)$, denote by ϕ_F the polynomial*

$$\phi_F(x) = \sum_{\alpha \in F \cap \mathbb{N}^d} c_\alpha x^\alpha.$$

We can now define the nondegeneracy condition we impose on our phase:

Definition 4 (Nondegeneracy)**.** *We say that an analytic function ϕ is* ***nondegenerate*** *if for all compact faces $F \subset N(\phi)$ the polynomials ϕ_F satisfy*

$$\|x \nabla \phi_F(x)\| \stackrel{def}{=} \max_{1 \leq i \leq d} |x_i \partial_i \phi_F(x)| \neq 0$$

for all x such that $x_1 x_2 \cdots x_d \neq 0$. [7]

[7] The algebraic variety (zero set) $\{x \in \mathbb{R}^d : x_1 \cdots x_d = 0\}$ is not the most general one we can

Nondegeneracy is equivalent to the property that for all x and all compact $F \subset N(\phi)$ there is some component of $\nabla\phi_F(x)$ that is nonzero away from the coordinate hyperplanes; the phrasing used in the definition is preferred because working with $x\nabla\phi_F(x)$ $(x\nabla\phi(x))$ is easier than working with $\nabla\phi_F(x)(\nabla\phi(x))$.[8] This convention makes the integration by parts argument in Chapter 5 much more natural.

Observation 1. *The fact that $N(\phi)$ has finitely many extreme points, together with the nondegeneracy condition, tells us there is a finite set of monomials that determines the behavior of $x\nabla\phi(x)$ for any x near the origin. Therefore for nondegenerate analytic functions, we only need information about a Taylor polynomial of some degree to recover $N(\phi)$ and all of its data. Namely, let $m = \max_{\beta \in N(\phi)}\{|\beta|\}$, where the maximum is taken over all extreme points. Then, since $\phi(x)$ is real analytic in a neighborhood of the origin, Taylor's theorem guarantees that for all m the function ϕ can be expressed as a* ***Taylor expansion of order m***

$$\phi(x) = \sum_{|\alpha| \le m} c_\alpha x^\alpha + \sum_{|\alpha| = m} h_\alpha(x) x^\alpha = P_m(x) + R_m(x), \tag{3.2.1}$$

where h_α is continuous and approaches 0 as $x \to \mathbf{0}$. Moreover, $P_m(x)$ is a unique polynomial of degree at most m and we call $R_m(x)$ the remainder. We see that $N(\phi) =$

consider for nondegeneracy. It is possible to work with any normal crossings singularity at the origin (intersection of at most d hyperplanes in $\mathbb{R}^d$), or more generally, that there is some nice change of variables so that the variety can be transformed into an arbitrary finite intersection of hyperplanes near the origin. One goal of future work is to consider singularities that break up into intersections of hyperplanes when perturbed. We do not say more about these conditions, as there is nothing in this document that proves anything about them.

[8] For example, see **P2** below.

$N(P_m)$ because every extreme point lies on or below the plane $\langle \xi, \mathbf{1} \rangle = m$. Moreover, we see that close enough to the origin each $|h_\alpha(x)| \lesssim \max_{1\le i\le d}|x_i|$ since h_α is analytic. Therefore $h_\alpha(x)x^\alpha$ cannot contribute most for any α.

Observation 1 is made rigorous in the proof of Lemma 1 when we see that $N(P_m)$ indeed contains all of the information we need about the decay of analytic functions.[9] We choose to consider this form of the remainder because it gives us more information about ϕ in the sense we care about.

From now on, when we decompose an analytic function by Taylor's theorem, namely $\phi = P_m + R_m$, we assume m is the smallest such that $N(\phi) = N(P_m)$. The same holds for C^k functions.

3.2.1 Analogous statements for convenient C^k functions.

Now assume that $\phi : \mathbb{R}^d \to \mathbb{R}$ is in the class C^k in a neighborhood of the origin. By Taylor's theorem, there are continuous $h_\alpha(x) \to 0$ as $x \to \mathbf{0}$ as in Observation 1 such that for all $m \le k$ we can write

$$\phi(x) = \sum_{|\alpha|\le m} c_\alpha x^\alpha + \sum_{|\alpha|=m} h_\alpha(x)x^\alpha = P_m(x) + R_m(x). \tag{3.2.2}$$

Using the end of the last section as motivation, we define the Newton polyhedron for ϕ that have small remainder in some sense. We need a condition that guarantees

[9] This is some type of compactness argument, as we first require monomials of every degree in the expansion of ϕ to know m, i.e., infinite information is being reduced to finite information.

the remainder can never contribute most near the origin, e.g., one cannot have P_m behaving like y^2 and R_m behaving like x^2; since we cannot quantify the decay of h, it could potentially behave like x, and that is bad. For simplicity, we assume that $N(P_m)$ intersects each coordinate axis if ϕ is not analytic. Such functions are well-studied and have certain nice properties related to decay near the origin. [10] Below, we define the Newton polyhedron of a C^k function satisfying this condition and then show it is well-defined. **Whenever ϕ is C^k for $k \geq 1$, we assume $\phi(\mathbf{0}) = 0$. Keep in mind we only care about functions satisfying $\nabla\phi(\mathbf{0}) = 0$ in this thesis.** To be consistent with recent literature, we define:

Definition 5 (Convenient function)**.** *Assume ϕ is C^k or smooth. If in some neighborhood of the origin we can write $\phi = P_m + R_m$ as in 3.2.2 and $N(P_m)$ intersects each coordinate axis, we say ϕ is $m-$**convenient**.*

Although in the literature a function is simply called **convenient** if its Newton polyhedron intersects each axis, for some of our results it is important to keep track of how we decompose ϕ, as many of our estimates depend on m. For example, the proof of Lemma 1 requires fixing a Taylor expansion $P_m + R_m$ for the phase under consideration. When the statements do not depend on m, we simply say ϕ is **convenient**.

Definition 6 (Taylor support of a convenient function)**.** *If ϕ is $m-$convenient, we define* $\text{supp}(\phi) = \text{supp}(P_m)$.

[10] As mentioned in the introduction, see for example [9], [7], and [19].

Definition 7 (Newton polyhedron of an $m-$convenient function)**.** *Let ϕ be $m-$ convenient for some $1 \leq m < \infty$. Writing $\phi = P_m + R_m$, we define $N(\phi) = N(P_m)$.*

$N(\phi)$ is well-defined because if ϕ has additional smoothness, the Newton polyhedron $N(P_m)$ contains $N(P_n)$ for all $n \geq m$ since $N(P_m)$ contains all vectors v satisfying $|v| \geq m$. With the Newton polyhedron defined for $m-$convenient phases, we can define the analogous nondegeneracy condition:

Definition 8 ($m-$convenient nondegeneracy)**.** *We say an $m-$convenient function $\phi = P_m + R_m$ is* ***nondegenerate*** *if P_m is nondegenerate.*

We now state a key properties of the functions under consideration: **P2** for analytic and $m-$convenient phases; property **P2** is particularly important for the proof of Lemma 2. We let ϕ be real analytic or else $m-$convenient and write $\phi = P_m + R_m$ for the rest of the section.

P1 For all x small enough, there is some $\beta = \beta_x \in N(\phi)$ satisfying $|\beta| = m$ such that $|x^\gamma \partial^\gamma R_m(x)| \lesssim x^\beta$.

P2 Let $\gamma \in \mathbb{N}^d$. If $|\gamma| \leq m$ then for all x small enough, $|x^\gamma \partial^\gamma \phi(x)| \lesssim x^\beta$ for some $\beta = \beta_x \in N(\phi)$, where the implicit constant is independent of x.

Proof. Property **P1** can be shown easily: apply Taylor's theorem to the function $\partial^\gamma \phi$: there are continuous $h_{m,\gamma}(x) \to 0$ as $x \to \mathbf{0}$ such that

$$|x^\gamma \partial^\gamma R_\alpha| = \Big| \sum_{|\alpha| = m - |\gamma|} h_{\alpha,\gamma} x^{\alpha+\gamma} \Big| \lesssim \max_{|\alpha| = m - \gamma} x^{\alpha+\gamma}.$$

The implicit constant depends on the functions $h_{m,\gamma}$. Since all $|\beta| = m$ lie in $N(\phi)$, so does $\alpha + \gamma$.

Now we prove **P2** for analytic functions. It is enough to show $\mathrm{supp}(\cdot^{\gamma}\partial^{\gamma}\phi)$ is contained in $\mathrm{supp}(\phi)$: for any analytic ϕ on $[-4,4]^d$, we know

$$\sup_{||x||_\infty \leq 1} |\phi(x)| \lesssim \max_{\alpha \in \mathrm{supp}(\phi)} x^{\alpha}$$

by rewriting ϕ as a polynomial plus remainder and applying **P1**. The implicit constant depends on the remainder term and the coefficients of the m^{th} order Taylor polynomial of ϕ. Therefore, since $\cdot^{\gamma}\partial^{\gamma}\phi$ is analytic,

$$\sup_{||x||_\infty \leq 1} |x^{\gamma}\partial^{\gamma}\phi(x)| \lesssim \max_{\alpha \in \mathrm{supp}(\cdot^{\gamma}\partial^{\gamma}\phi)} x^{\alpha}. \tag{3.2.3}$$

Indeed, we have the containment we seek: given any monomial x^{α} with $\alpha \in \mathrm{supp}(\phi)$, observe that $x^{\gamma}\partial^{\gamma}x^{\alpha} = cx^{\alpha}$ for some $c \in \mathbb{R}$. Analyzing the cases where $c = 0$ and where $c \neq 0$, we see that $\mathrm{supp}(\cdot^{\gamma}\partial^{\gamma}\phi) \subset \mathrm{supp}(\phi)$, and therefore the maximum over $\mathrm{supp}(\cdot^{\gamma}\partial^{\gamma}\phi)$ in (3.2.3) is bounded above by the maximum over $\mathrm{supp}(\phi)$ and the claim is finished.

Next, if ϕ is $m-$convenient, write $\phi = P_m + R_m$ and apply **P1** together with the bound on the analytic function P_m proven in the preceding paragraph. □

In the future we abuse notation and write $P_{m,\gamma}$ and $h_{m,\gamma}$, respectively, as the right side of

$$\sum_{|\alpha|\le m-|\gamma|} c_{\alpha,\gamma}x^{\alpha+\gamma} + \sum_{|\alpha|=m-|\gamma|} h_{\alpha,\gamma}x^{\alpha+\gamma} = \sum_{|\alpha|\le m} c_{\alpha,\gamma}x^{\alpha} + \sum_{|\alpha|=m} h_{\alpha,\gamma}x^{\alpha}. \tag{3.2.4}$$

3.3 Normal vectors of the Newton polyhedron

We briefly discuss a subset of linear functionals on $\mathbb{R}^d$, namely the set

$$\{w \in \mathbb{R}^d : \langle \xi, w\rangle = 1 \text{ for some } \xi \in \partial N(\phi)\}.$$

Although it is possible to think about this set as a polyhedron (the dual polyhedron), we think of the set above as just the set of normal vectors to supporting hyperplanes $H \not\ni \mathbf{0}$ without any geometric structure. In particular, we care most about the finitely many normals to codimension 1 faces not contained in coordinate hyperplanes.

To introduce Theorem 2, we need to introduce a convention for normal vectors to supporting hyperplanes of $N(\phi)$ not containing the origin. From now on, when we talk about supporting hyperplanes H of $N(\phi)$ we mean *only* those not containing the origin, and we use a normalization convention for normal vectors of such hyperplanes H: we pick the unique vector $w \in \mathbb{R}^d_{\geq}$ satisfying $H = \{\xi : \langle \xi, w\rangle = 1\}$. Any supporting hyperplane H of $N(\phi)$[11] can be defined this way, and we write H_w for such hyperplanes, namely $H_w = \{\xi : \langle \xi, w\rangle = 1\}$. There is a nice geometric way to

[11]It is very important to remember that $\mathbf{0} \notin H$. In particular, supporting hyperplanes of $N(\phi)$ containing the origin, i.e., those containing coordinate hyperplanes, are never considered.

show such normals exist: if $w = (w_1, \ldots, w_d)$, then H_w intersects the coordinate axes at $x_i = w_i^{-1}$, whenever $w_i \neq 0$ (and does not intersect the x_i axis if $w_i = 0$). Such normals are guaranteed to exist since $\phi(\mathbf{0}) = 0$ and ϕ not identically zero implies $N(\phi)$ does not contain the origin. Thinking about normals geometrically this way, we can easily see that the normal w has some $w_j = 0$ if and only if w is a normal to a hyperplane intersecting $N(\phi)$ in some unbounded face. For example, if the face contains α then it must contain $\alpha + n\mathbf{e}_j$ for all $n \in \mathbb{N}$. In this document, unbounded codimension 1 faces $F \subset N(\phi)$ exist only if ϕ is not convenient, i.e., only for analytic ϕ.

We say w is a **corresponding normal** of the face F of $N(\phi)$ if $H_w \cap N(\phi) = F$. Note that we can say *the* corresponding normal of F if F is codimension 1. We also say w is a normal of $N(\phi)$ if w is a normal of any $F \subset N(\phi)$.

The condition originally assumed by Varchenko in [18] to estimate $I(\lambda)$ was that the phase ϕ is nondegenerate. In order to make analogies with his estimates, we need the following definition.

Definition 9 (Newton distance). *Let ϕ be real analytic. The **Newton distance** of ϕ is the infimum*

$$t = \inf\{s \in (0, \infty) : (s, \ldots, s) \in N(\phi)\}.$$

If ϕ is $m-$convenient, the Newton distance is defined to be the Newton distance of P_m.

In this document, the variable t is always reserved for the Newton distance of the phase under consideration, and in particular, $\mathbf{t}$ is always an element of $N(\phi)$.

One can check $t \geq 1/d$ if $\phi(\mathbf{0}) = 0$ and not identically zero: if $\operatorname{supp}(\phi)$ contains each standard unit coordinate vector $\mathbf{e}_i$, then $t = 1/d$ since the vector $\mathbf{1/d}$ can be expressed as the convex combination $\sum_{i=1}^{d} \mathbf{e}_i/d$. By definition, the Newton polyhedron of any analytic function is contained in the Newton polyhedron of $x_1 + \cdots + x_d$, so $t \geq 1/d$.

If some scalar multiple of β lies in $N(\phi)$,[12] we use the convention of writing $w(\beta)$ for the set $\{w : \langle \beta, w \rangle \text{ is minimal}\}$ where the minimum is taken over all finitely many normals w of codimension 1 faces F of $N(\phi)$; which polyhedron we are talking about is always clear from context. In particular, $\langle \beta, w^1 \rangle = \langle \beta, w^2 \rangle$ for all $w^1, w^2 \in w(\beta)$. Note that $k_j = \min\{|w(\beta)|, d\}$ is equal to the highest codimension over all faces containing some multiple of β. The minimum is required because it is possible for a vertex to be written as an intersection of more than d faces. In particular $|w(\beta)| \leq d$ implies $k_j = |w(\beta)|$. If $\gamma \in \mathbb{R}^d_{\geq}$, we define $\langle \gamma, w(\beta) \rangle$ to be the scalar $\max_{w \in w(\beta)} \langle \gamma, w \rangle$. This convention is used mainly in Chapter 7 and is important for determining when there are no multiples of β and γ lying in the same codimension 1 faces: in the case that H_w contains β but not γ, we have

$$\langle \gamma, w(\beta) \rangle \geq \langle \gamma, w \rangle > \langle \beta, w \rangle = \langle \beta, w(\beta) \rangle.$$

[12] For all $\alpha \in \mathbb{N}^d$, some multiple of $\alpha + \mathbf{1}$ is guaranteed to lie in $N(\phi)$. For example, $\langle t(\alpha + \mathbf{1}), w \rangle \geq \langle \mathbf{t}, w \rangle \geq 1$ for all normals w of $N(\phi)$.

Observation 2. *For any real number $c > 0$, we can show*

1. $\langle \gamma, w(\beta)\rangle = \langle c\gamma, w(\beta)\rangle / c$, *and*

2. $\langle \gamma, w(\beta)\rangle = \langle \gamma, w(c\beta)\rangle$.

To see both of these facts, simply use the definition of $\langle \gamma, w(\beta)\rangle$:

$$\max_w \langle c\gamma, w\rangle = c \max_w \langle \gamma, w\rangle.$$

Assume we are given a Newton polyhedron with Newton distance t. Then $\mathbf{t}$ lies on a face of the polyhedron whose supporting hyperplanes cannot contain the origin, so $\langle \mathbf{t}, w(\mathbf{t})\rangle = 1$. Therefore, by Observation 2, $\langle \mathbf{1}, w(\mathbf{1})\rangle = 1/t$. This equality is an important bridge with respect to Varchenko's result and this document: first notice $\langle \mathbf{1}, w(\mathbf{1})\rangle \leq \langle \beta + \mathbf{1}, w(\beta + \mathbf{1})\rangle$ for all $\beta \in \mathbb{N}^d$. For any $w \in w(\beta + \mathbf{1})$, the inner product $\langle \beta + \mathbf{1}, w\rangle$ equals $\langle \mathbf{1}, w\rangle + \langle \beta, w\rangle \geq 1/t + \langle \beta, w\rangle$. The remaining quantity can be bounded above by $1/t$ since β and w have nonnegative entries. It is no coincidence that $\langle \mathbf{1}, w(\mathbf{1})\rangle$, the smallest such inner product, is the exponent of the first term appearing in the asymptotic expansion of $I(\lambda)$, as shown by Varchenko for analytic phases with $t > 1$. We continue this discussion after the statement of Theorem 2.

Finally, we conclude this section by mentioning that if $F = H_w \cap N(\phi)$ has dimension k, then F contains $k + 1$ linearly independent vectors. This is because all H_w under consideration cannot contain the origin, and therefore the $k + 1$ affinely independent points in F must be linearly independent. This is just by definition: assume $\sum_{i=1}^k \lambda^i v^i = \mathbf{0}$ and assume without loss of generality $\lambda_0 \neq 0$. Then $\sum_{i=0}^k \lambda_i v^i = \mathbf{0} \iff \sum_{i=1}^k \lambda_i v^i = -\lambda_0 v^0 \neq \mathbf{0}$.

3.4 Preliminary geometric results about $w(\beta)$.

We now prove the following facts for use in Chapter 7:

Proposition 2. *Let ϕ be analytic and assume $N(\phi)$ does not contain the origin. Let $\beta \in \mathbb{N}^d$ have all positive components. In particular, we can define $w(\beta)$ and $w(\alpha+\beta)$ for any $\alpha \in \mathbb{N}^d$. Let $w \in w(\beta)$.*

If $\langle \beta, w\rangle = \langle \alpha+\beta, w(\alpha+\beta)\rangle$, then

P3 *$w(\alpha+\beta) \subseteq w(\beta)$ and $\alpha_j w_j = 0$ for all $w \in w(\alpha+\beta)$,*

P4 *either $w(\alpha+\beta) \subsetneq w(\beta)$ or else $\alpha \neq \mathbf{0}$ implies β does not lie in any compact codimension 1 face, and*

P5 *If $\alpha \neq \mathbf{0}$, then $|w(\alpha+\beta)| \leq |\{j : \alpha_j \neq 0\}| \leq d-1$.*

P6 *If instead we assume $\alpha \in N(\phi)$ and $\langle \beta, w\rangle + 1 = \langle \alpha+\beta, w(\alpha+\beta)\rangle$, then $|w(\alpha+\beta)| \leq d$ with equality only if $\alpha = \beta$.*

Proof. We first prove **P3**: for all $w' \in w(\alpha+\beta)$ we have

$$\langle \beta, w\rangle = \langle \alpha+\beta, w(\alpha+\beta)\rangle = \langle \alpha+\beta, w'\rangle \geq \langle \beta, w\rangle + \langle \alpha, w'\rangle,$$

so we conclude $\langle \alpha, w(\alpha+\beta)\rangle = 0$. Therefore $\langle \beta, w\rangle$ equals

$$\langle \beta, w(\alpha+\beta)\rangle = \max_{w'' \in w(\alpha+\beta)} \langle \beta, w''\rangle \geq \min_{w'' \in w(\alpha+\beta)} \langle \beta, w''\rangle \geq \min_{w'' \in w(\beta)} \langle \beta, w''\rangle.$$

Since the last term also equals $\langle \beta, w(\beta)\rangle$, this implies $w(\alpha+\beta) \subseteq w(\beta)$ since $w(\beta)$ contains all normals to codimension 1 faces satisfying $\langle \beta, w\rangle = \langle \beta, w(\beta)\rangle$.

Next, we prove **P4**: assume $w(\alpha+\beta) = w(\beta)$ for some $\alpha \neq \mathbf{0}$. Property **P3** implies there is some $1 \leq j \leq d$ such that $w_j = 0$ for all $w \in w(\alpha+\beta) = w(\beta)$. Therefore, since each normal corresponding to a compact codimension 1 face must have positive components, β does not lie in any compact codimension 1 face.

To show **P5**, the most important result we need is **P3**: $\alpha_j w_j = 0$ for all $w \in w(\alpha+\beta)$. For brevity, write L for the line $\{s(\beta+\mathbf{1}) : s \in \mathbb{R}\}$. Let $v \in L \cap N(\phi)$. Without loss of generality, assume $\alpha_1, \ldots, \alpha_k \neq 0$ and the rest of the components of α are zero. Any $d-k+1 \leq d$ vectors in $w(\alpha+\beta)$ with exactly k zero components must be linearly dependent. Assume without loss of generality that $\{w^1, \ldots, w^{d-k}\} \subseteq w(\alpha+\beta)$ is a maximal linearly independent set (it could contain less than $d-k$ vectors). Let $H = \bigcap_{i=1}^{d-k} H_{w^i}$. The set H is nonempty because any intersection of hyperplanes with normals in $w(\alpha+\beta)$ contains v. Any $w \in w(\alpha+\beta)$ can be written as a linear combination $w = \sum a_i w^i$ and in fact $\sum a_i = 1$ since $\langle v, w^i\rangle = \langle v, w\rangle = 1$ for all $1 \leq i \leq d-k$. Therefore $H_w \cap H = H$, as the restriction $\langle \xi, w\rangle = 1$ adds no additional information than what was already encompassed in H, so it cannot be a normal to a codimension 1 face: the intersection of any $d-k+1$ distinct codimension 1 faces must be at most codimension $d-k+1$ for $1 \le k \le d$ (see [21, Chapter 2]). Since the intersection H is nonempty, it must be codimension at most $d-k$. We conclude $|w(\alpha+\beta)| \le d-k$.

Finally, assume $\alpha \in N(\phi)$ and $\langle \beta, w \rangle + 1 = \langle \alpha+\beta, w(\alpha+\beta) \rangle$. Then there is some w' normal to a codimension 1 face such that

$$\langle \beta, w \rangle + 1 = \langle \alpha+\beta, w' \rangle = \langle \alpha, w' \rangle + \langle \beta, w' \rangle \ge 1 + \langle \beta, w' \rangle \ge \langle \beta, w \rangle + 1.$$

Therefore $\langle \beta, w' \rangle = \langle \beta, w \rangle = \max_{w \in w(\beta)} \langle \beta, w \rangle$, so $w' \in w(\beta)$. Similarly, since $\langle \alpha, w' \rangle = 1$, we must have $w' \in w(\alpha)$. Therefore $w(\alpha+\beta) \subseteq w(\alpha) \cap w(\beta)$. Since the nonempty intersection of d hyperplanes is a vertex, if $w(\alpha)$ and $w(\beta)$ share at least d elements $w^1, \ldots, w^d$ in common, we must have $\alpha = \bigcap_{i=1}^{d} H_{w^i} = \beta$. $\square$

These properties are required to prove the asymptotic expansion of $I(\lambda)$.

3.5 Back to Theorem 2

3.5.1 The set $\mathcal{C}$ of exponents

The asymptotic expansion of $I(\lambda)$ of Theorem 2 has exponents that are easy to define with the help of the Newton polyhedron, and many are easy to visualize geometrically. However, there are some exponents that are harder to visualize. These exponents are of the form $n - \langle \beta, w(\beta) \rangle$ for some $n \in \mathbb{N}$ and some $\beta = \alpha^1 + \cdots + \alpha^n + \mathbf{1}$, where each $\alpha^i \in N(\phi)$.[13] The ones not of the form $-\langle \beta, w(\beta) \rangle$ come from elements $\alpha + \mathbf{1}$ for some $\alpha \in \operatorname{supp}(\phi)$ that cannot be decomposed as $u + v$ for some $u \in \partial N(\phi) \cap \mathbb{N}^d$ and some $v = \alpha + \mathbf{1}$. We now show this occurs naturally.

Example 1. *Consider the* $5-$*convenient phase*

$$\phi(x, y) = x^4 + y^5 + x^2 y^3.$$

The Newton polyhedron contains everything in the first quadrant lying above the line $5x + 4y = 20$. *This line represents a compact codimension 1 face with normal* $w = (1/4, 1/5)$. *The point* $(2, 3)$ *lies above this face, since* $\langle (2, 3), w \rangle = 11/10 > 1$. *Also,* $(2, 3) + (1, 1) = (3, 4)$ *is such that* $\langle (3, 4), w \rangle = 31/20$. *In this case,* $\alpha = (2, 3)$, *so* $\alpha + \mathbf{1} = (3, 4)$. *The vector* $(3, 4)$ *cannot be decomposed as* $u + v$ *for* $u \in \partial N(\phi) \cap \mathbb{N}^d$ *and* $v \in \mathbb{N}^d$.

Such points prevent us from writing all the exponents in Theorem 2 as $-\langle \beta, w(\beta) \rangle$ *because to get higher order terms, we need to estimate* $\lambda \frac{d}{d\lambda} I(\lambda)$. *In this case, we can split* $\lambda \frac{d}{d\lambda} I(\lambda)$ *into two integrals we can easily estimate with exponent of the form* $-\langle \beta, w(\beta) \rangle$, *plus one more:*

$$i\lambda \int e^{i\lambda\phi(x)} x^2 y^3 \psi(x) dx.$$

[13] If, for example, the amplitude ψ satisfies certain decay properties, all exponents are of the form $-\langle \beta, w(\beta) \rangle$. However, this thesis is concerned with the general expansion for arbitrary amplitudes. To learn more about how the decay of ψ affects the exponents, see [9].

Theorem 1 then guarantees we can bound this integral by $\lambda^{1-\langle(3,4),(1/4,1/5)\rangle}$. *We know* $1-\langle(3,4),(1/4,1/5)\rangle < -1/t$. *However,* $1-\langle(3,4),(1/4,1/5)\rangle \neq -\langle\beta, w(\beta)\rangle$ *for any* $\beta \in \mathbb{N}^d$ *because* $11/20 \neq a/4+b/5$ *for any positive integers* a *and* b.

Almost by definition, we know that all these bad points lie inside

$$D = N(\phi) - \bigcup_{v\in\partial N(\phi)\cap\mathbb{N}^d} v+\mathbb{R}^d_{\geq},$$

so there are few of them in some sense. If $D\cap\mathbb{N}^d = \varnothing$, we have a much more convenient way to express the exponents appearing in Theorem 2.

Using induction and the results built above, we are able to prove Theorem 2. With respect to each phase function ϕ, we let $w(\alpha+\mathbf{1})$ be as in Section 3.3 for $\alpha\in\mathbb{N}^d$. Recall that $|w(\alpha+\mathbf{1})|$ is equal to the number of codimension 1 faces containing $\alpha+\mathbf{1}$. Therefore the largest codimension of any face containing alpha is equal to $\min\{d, |w(\alpha+\mathbf{1})|\}$; the minimum is necessary because it is possible for a vertex to lie in any number of codimension 1 faces. When $|w(\alpha+\mathbf{1})|\leq d$, indeed $|w(\alpha+\mathbf{1})|$ is equal to the largest codimension over all faces containing α. Let $\mathcal{C} =$

$$\{\langle\beta, w(\beta)\rangle - n : \beta = \alpha^1+\cdots+\alpha^n+\mathbf{1}\in\mathbb{N}^d \text{ for some } n\in\mathbb{N} \text{ and } \alpha^i\in N(\phi)\}.$$

3.5.2 Clarifications about Theorem 2

Some statements of Theorem 2 require further clarification. First, we show that $\{\langle\alpha+\mathbf{1}, w(\alpha+\mathbf{1})\rangle\}_{\alpha\in\mathbb{N}^d}$ runs through finitely many arithmetic progressions of positive rationals. Each normal w of a codimension 1 face F of $N(\phi)$ can be uniquely defined by d linearly independent vectors α^i in $\text{supp}(\phi)\cap F$. If A is the matrix with rows α^i, then w must satisfy $Aw = \mathbf{1}$, by definition. Hence, $w = A^{-1}\mathbf{1}$. The matrix A^{-1} must have rational entries, since A has rational entries, therefore $w \in \mathbb{Q}^d$. In fact, each component of w must be nonnegative because it is oriented towards the interior of the Newton polyhedron, so in particular it must be in $\mathbb{R}^d_{\geq}$. The Newton polyhedron has finitely many codimension 1 faces, so there are finitely many such w each with d components w_j. Writing each $w_j = r_j/q_j$ over a fixed normal w where r_j, q_j are integers, let q_w be the lowest common multiple of the q_j. The arithmetic progressions come from the rationals $1/q_w$ over all normals w to codimension 1 faces of $N(\phi)$. More generally, if $c \in \mathcal{C}$ then $c = \langle\beta, w(\beta)\rangle - n > 0$ for some n. Writing c as a fraction, it is also easy to see that $c = r/q_w$ for some $r \in \mathbb{N}$ and $w \in w(\beta)$.[14]

Varchenko showed that the first term of the expansion (2.2.3) with nonzero coefficient is $\lambda^{-1/t}\log^{d_{\mathbf{t}}-1}(\lambda)$, where $d_{\mathbf{t}}$ is the largest codimension over all faces containing $\mathbf{t}$, assuming the Newton distance of ϕ is larger than 1. Indeed, this is the first term guaranteed by Theorem 2 (by the discussion following Observation 2). Also by Observation 2, we see that there is an easy geometric way to describe the exponents in Theorem 2

[14] Going back to Example 1, $31/20-1 = 11/20$. Here, the lowest common multiple is $q = 4\cdot 5 = 20$.

without parametrizing the Newton polyhedron. With our parametrization, the normals are such that $c = \langle \alpha, w(\alpha) \rangle$ is the scaling required so that $\alpha/c \in F = H_w \cap N(\phi)$ for any $w \in w(\alpha)$.

3.5.3 Examples

> *As the use of the [polyhedron] is better shown by examples than by description...*- Sir Isaac Newton[15]

We try to illustrate how exactly to get the first few terms of the asymptotic expansion of an oscillatory integral via simple geometric considerations.

Example 1: Let $\phi(x, y, z) = x^8 + y^3 - z^2$. The analytic function ϕ is certainly nondegenerate. $N(\phi)$ has one codimension 1 face F (not lying in a coordinate hyperplane) with normal $w = (1/8, 1/3, 1/2)$. Since F intersects each coordinate axis, every line containing both the origin and $\beta + \mathbf{1}$ passes through F for any $\beta \in \mathbb{N}^d$. In particular, the j^{th} power of log in the expansion is $d_j = 0$ for all j. Theorem 2 tells us that for smooth ψ supported close enough to the origin,

$$I(\lambda) \sim \sum_{j=0}^{\infty} \lambda^{-p_j},$$

where $p_0 < p_1 < \cdots$ is the ordering of $\{\langle \beta + \mathbf{1}, w(\beta + \mathbf{1}) \rangle : \beta \in \mathbb{N}^d\}$. Here the p_j contain *all* arithmetic progressions in $1/8, 1/3, 1/2$, since β is arbitrary and there is

[15]The Newton polyhedron originated in Newton's *Enumeration of Lines of the Third Order, Generation of Curves by Shadows, Organic Description of Curves, and Construction of Equations by Curves*. A translation can be found in [17].

only one normal vector. In fact that Theorem 2 says we need to consider every $n/24$ for $n > 23$. Letting $\beta = \mathbf{0}$, we see that

$$p_0 = 1/8 + 1/3 + 1/2 = 23/24.$$

Letting $\beta = (3, 2, 0)$ and $n = 1$ gives $4/8 + 3/3 + 1/2 - 1 = 1$, the second highest exponent. Next, $p_2 = 3/8 + 5/3 - 1 = 25/24$, seen by taking $\beta = (2, 4, 0)$.Letting $\beta = (9, 0, 0)$ give $p_3 = 10/8 + 1/3 + 1/2 - 1 = 26/24$.

Example 2: Let $\phi = P_8 + R_8$ be any C^k function such that $P_8 = x^8 + y^3 - z^2$, a nondegenerate $8-$convenient polynomial. For k large enough, the theorem tells us

$$I(\lambda) \sim \sum_{j=0}^{2} a_j(\psi)\lambda^{-p_j},$$

where p_j are ordered as in the previous example. Since $p_3 = 26/24$, we also know

$$\left| I(\lambda) - \sum_{j=0}^{2} a_j(\psi)\lambda^{-p_j} \right| \lesssim \lambda^{-26/24}.$$

Example 3: Let $\phi(x, y) = y^5 - xy^3 + x^3y - x^4y$. There are three codimension 1 faces (not contained in a coordinate hyperplane) defined by vertices $(0, 5)$ and $(1, 3)$, $(1, 3)$ and $(3, 1)$, and an unbounded face generated by $(3, 1) + s(1, 0)$. We can check ϕ is nondegenerate. Corresponding to these faces are the three normals

$$w^1 = (2/5, 1/5), w^2 = (1/4, 1/4), w^3 = (0, 1).$$

Since $|w^2| = 1/2 < |w^1| = 3/5 < |w^3| = 1$, we see $p_0 = 1/2$. $\beta = (1, 0)$ and $(0, 1)$ give us

$$\langle \beta + \mathbf{1}, w(\beta + \mathbf{1}) \rangle = \langle \beta + \mathbf{1}, w^2 \rangle = 3/4.$$

However, $\beta = (0, 5)$ gives $\langle \beta + \mathbf{1}, w^1 \rangle - 1 = 3/5$. Indeed $p_1 = 3/5$ and $p_2 = 3/4$. Then, we can take $\beta = (1, 2)$ to get $\langle \beta + \mathbf{1}, w^1 \rangle - 1 = 4/5$. This β does not give us a log term and neither does $\beta = (2, 2)$. However,

$$\langle (1, 3), w^1 \rangle = \langle (1, 3), w^2 \rangle = \langle (3, 1), w^2 \rangle = \langle (3, 1), w^3 \rangle = 1.$$

Therefore together with $p_3 = 1$, we have a nonzero exponent of $\log$. In this case, the integral $I(\lambda)$ behaves like

$$I(\lambda) \sim \lambda^{-1/2} + \lambda^{-3/5} + \lambda^{-3/4} + \lambda^{-4/5} + \lambda^{-1} \log(\lambda) + \lambda^{-1} + \cdots$$

Here the exponents can only be of the form $a/5$ or $b/4$ for arbitrary $a, b \in \mathbb{Z}_+$. When these two cannot be equal, there is no log term paired with the exponent of λ; we used this fact for $p_1 = 3/5$, $p_2 = 3/4$ and $p_3 = 4/5$. Otherwise, there is a log term. We conclude that exponents are of the form $n, n + 1/5, n + 1/4, n + 2/5, n + 1/2, n + 3/5, n+3/4$, and $n+4/5$; log terms only appear with integer exponents of λ. Moreover, the exponents p_j are all of the form $\langle \beta + \mathbf{1}, w(\beta + \mathbf{1}) \rangle$ except $p_3 = 4/5$.

Chapter 4

Proof of Lemma 1

4.1 Motivation

Our motivation for Lemma 1 is for the proof of Lemma 2: we integrate the left side of (2.2.1) by parts N times. Lemma 1 is certainly interesting in its own right, reminiscent of Lojasiewicz's famous theorem [14, Theorem 17] (an English version can be found in [13]). Unfortunately, Lojasiewicz's theorem does not imply the result we are looking for in the proof of Lemma 2, even for analytic functions: Lemma 1 has stronger assumptions, but gives a much stronger result. Greenblatt also proved a very nice version of Lemma 1 in [7], namely Lemma 3.6, under an assumption on the order of the zero of analytic functions but not on the zero set. It worked well in his setting, but unfortunately does not work in ours since we do not have an assumption on the order of the singularity.

From now on, x always lies in $[1,4]^d$ and we scale by ε when talking about elements outside the box $[1,4]^d$. This chapter does not involve any integrals, so we use $i \in \mathbb{N}$ as an index.

We choose to parametrize ε close to the origin by normals of $N(\phi)$ for the rest of the chapter in order to better visualize how exactly the normals of the Newton polyhedron determine which elements in $\mathrm{supp}(\phi)$ are the largest. We first prove that it is possible to parametrize elements of $\mathbb{R}^d$ close enough to the origin by normals of supporting hyperplanes.

Let $0 < \tau < 1$. We show for all $\varepsilon \in (0,\tau)^d$ there is some $S \in (0, \tau^{1/d})$ and some supporting hyperplane H_w of $N(\phi)$ such that $S^w = \varepsilon$, and therefore $S = S^{\langle \alpha, w\rangle} = (S^w)^\alpha = \varepsilon^\alpha$ for all $\alpha \in H_w$, and in particular $\alpha \in H_w \cap \mathrm{supp}(\phi)$.

First note that the $d-$tuple $(1/d, \ldots, 1/d)$ lies on or below $N(\phi)$. Therefore, for all $\alpha \in N(\phi)$ there is some $1 \leq i \leq d$ such that $\alpha_i \geq 1/d$. If H_w is a supporting hyperplane of $N(\phi)$ containing α, then for some $1 \leq i \leq d$ we can write

$$1 = \langle \alpha, w\rangle \geq \alpha_i w_i \geq w_i/d,$$

since every component of α and w are nonnegative. Hence, for every supporting hyperplane H_w there is some $1 \leq i \leq d$ such that $w_i \leq d$.

Next, let $\varepsilon \in (0,\tau)^d$. Assuming that ε_1 is the largest, we can solve the equations $\varepsilon_1^{q_i} = \varepsilon_i$ where $q_i \geq 1$. For all $q \in \mathbb{R}^d$ with positive components there is some supporting hyperplane H_w of $N(\phi)$ and positive constant c such that $q = cw$: we can just take a hyperplane with normal q, and translate in the direction of q (or $-q$) so that the hyperplane intersects only $\partial N(\phi)$. Since $q_1 = 1 \leq q_i$, we see that $w_1 \leq w_i$ and therefore $w_1 \leq d$. Hence, $1 = q_1 = cw_1 \leq cd$. Now we can solve for S in the required interval:

$$\varepsilon_i = \varepsilon_1^{q_i} = \varepsilon_1^{cw_i} = (\varepsilon_1^c)^{w_i},$$

so that $S = \varepsilon_1^c \leq \tau^c \leq \tau^{1/d}$. The first inequality holds because $\varepsilon_1 \leq \tau$ and the last inequality holds because $1 \leq cd$ and $0 < \tau < 1$. Define Hyp(ϕ)= $\{w \in (0,\infty)^d : H_w$ is a supporting hyperplane of $N(\phi)\}$. Since w and cw cannot both be normals to $N(\phi)$ for $c \neq 1$, to each ε corresponds a unique S and w such that $\varepsilon = S^w$, so this correspondence is a bijection between $(0,\tau)$ and some subset of $(0,\tau^{1/d})\times$Hyp(ϕ). Therefore we can just think of it as a reparameterization, and state this fact as a proposition:

Proposition 3 (Parameterization by supporting hyperplanes)**.** *Let* $\phi : \mathbb{R}^d \to \mathbb{R}$ *be real analytic in a neighborhood of the origin. Parametrize each supporting hyperplane* $H = H_w$ *of* $N(\phi)$ *by*

$$H_w = \{\langle \xi, w\rangle = 1 : \xi \in \mathbb{R}^d\}.$$

Let $0 < \tau < 1$*. There is a bijection between a subset of* $(0,\tau^{1/d})\times$*Hyp*(ϕ) *and* $(0,\tau)^d$ *with inverse defined by*

$$(S, w) \mapsto S^w \in (0,\tau)^d.$$

We now illustrate the ideas used to approach Lemma 1. Write $\phi = P_m + R_m$. By (3.2.4), for each $1 \leq i \leq d$ we can write $y_i \partial^{\mathbf{e}_i} \phi(y)$ as

$$\begin{aligned} y_i \partial^{\mathbf{e}_i} \phi(y) &= \sum_{|\alpha| \leq m} c_{\alpha, \mathbf{e}_i} y^\alpha + \sum_{|\alpha| = m} h_{\alpha, \mathbf{e}_i}(y) y^\alpha. \\ &= \sum_{|\alpha| \leq m} \alpha_i c_\alpha y^\alpha + \sum_{|\alpha| = m} h_{\alpha, \mathbf{e}_i}(y) y^\alpha \end{aligned} \tag{4.1.1}$$

For each compact $F \subset N(\phi)$, we can write the polynomial $y_i \partial^{\mathbf{e}_i} P_m(y)$ in (4.1.1) as

$$\sum_{\alpha \in F} c_\alpha \alpha_i y^\alpha + \sum_{\alpha \notin F} c_\alpha \alpha_i y^\alpha. \tag{4.1.2}$$

The left side of (4.1.2) equals the i^{th} component of $y \nabla \phi_F(y)$, which is a nonzero vector by nondegeneracy. The goal is to show for all y small enough there are F and $1 \leq i \leq d$ so that the significant contribution comes from some component of $y \nabla \phi_F(y)$. If for all $y = \varepsilon x \in [\varepsilon, 4\varepsilon]$ we can find $F \subset N(\phi)$ compact so that the main contribution comes from the sum over F, we are able to conclude that for some $1 \leq i \leq d$, (4.1.1) is bounded below by a uniform constant times

$$\left| \sum_{\alpha \in F} c_\alpha \alpha_i y^\alpha \right| = \left| \sum_{\alpha \in F} c_\alpha \alpha_i x^\alpha \varepsilon^\alpha \right| = \left| S \sum_{\alpha \in F} c_\alpha \alpha_i x^\alpha \right| \gtrsim S = (S^w)^\alpha \varepsilon^\alpha$$

for all $\alpha \in F = H_w \cap N(\phi)$, where $\varepsilon = S^w$. Indeed, we can show this by finding a compact face F so that the terms ε^α contribute most when $\alpha \in F$, and then we conclude (4.1.1) is bounded below by ε^α for all $\alpha \in N(\phi)$. The difficulty is in showing the second sum of (4.1.2) is negligible for appropriate F; it is much easier to show the remainder is small by **P1** and because $h_{\alpha,\mathbf{e}_i} \to 0$. We recursively define finitely many boxes $[b, 4b^{-1}]^d$, where $0 < b < 1$, on which we apply nondegeneracy, because the right side of (4.1.2) is not always negligible if we naively try to use the logic presented above. We might need to move to lower codimension faces $F_{d-1} \supseteq \cdots \supseteq F_0 = F$ by moving relatively large summands of the right-hand sum of (4.1.2) to the left-hand sum, checking whether all the summands in the right-hand side are negligible, and applying nondegeneracy on larger and larger boxes away from coordinate axes depending only on the polynomial P_m. During this process, we may also need to switch the partial derivative under consideration. This is the content of the main proposition below: Proposition 6.

4.1.1 Example

Let $\phi(y) = y_1^3 y_2 - y_2^2 = P_4(y)$. We can easily check that ϕ is nondegenerate: $y\nabla\phi(y) = (3y_1^3 y_2, y_1^3 y_2 - 2y_2^2) \neq \mathbf{0}$ for $y_1 y_2 \neq 0$ because the first component cannot be zero unless $y_1 y_2 = 0$. Let $\alpha^1 = (3, 1)$ and $\alpha^2 = (0, 2)$; these are the two extreme points of $N(\phi)$. The Newton polyhedron of ϕ has two vertices $F_1 = \{\alpha^1\}$ and $F_2 = \{\alpha^2\}$, a compact face F_3 that is equal to the convex hull of $\{\alpha^1, \alpha^2\}$, and the last face with supporting hyperplane not containing the origin is the set $F_4 = \{\alpha^1 + \ell(1, 0) : \ell \in \mathbb{R}_{\geq}\}$. We do not consider F_4 until the proof of Theorem 2 because unbounded faces correspond to some component of y vanishing.

Consider first the normal vector $w_3 = (1/6, 1/2)$ of the supporting hyperplane containing F_3. Notice that we chose w_3 so that $\langle \alpha^i, w \rangle = 1$ for $i = 1, 2$ because that is how we defined normal vectors to supporting hyperplanes. Then for $S > 0$ small, we have

$$(S^w x)\nabla\phi(S^w x) = (3(S^w x)^{\alpha^1}, (S^w x)^{\alpha^1} - 2(S^w x)^{\alpha^2}) = S(3x^{\alpha^1}, x^{\alpha^1} - 2x^{\alpha^2}).$$

In this case, since the hyperplane contains both vertices α^1 and α^2, the scaling is identical. If the scaling is very close to S^w, we choose the first component when it is larger, giving us a bound of $3Sx^{\alpha^1} \geq S$ for $x \in [1, 4]$. Otherwise, if the second component is much larger, we get the bound $S|(x^{\alpha^1} - 2x^{\alpha^2})| \gtrsim S$ because, loosely speaking, if the second component is large, it has to be far away from zero. This gives us estimates over boxes $[\varepsilon, 4\varepsilon] = [S^w, 4S^w]$. Finally, ε^α is the largest term since $S \geq S^{\langle \beta, w \rangle}$ for $\beta \in N(\phi)$ and therefore $\varepsilon^\alpha \geq \varepsilon^\beta$ so we can bound below by all ε^β. If the supporting hyperplane H_w contains α^1 and is far from α^2, then $\langle \alpha^2, w \rangle > 1 + \delta$, so $2(S^w x)^{\alpha^2}$ is very small if $\delta > 0$ is far enough from the origin, etc. The difficulty appears when δ is very small, and this is where we need to be careful– this is the most difficult obstruction in the proof of Lemma 1, and therefore, the most difficult part of bypassing resolution of singularities in our generalization of Varchenko's result.

In more detail, consider $(S^w x)^{\alpha^1} - 2(S^w x)^{\alpha^2} = -2Sx^{\alpha^2} + S^{\langle \alpha^1, w \rangle} x^{\alpha^1}$, when w is a normal to a supporting hyperplane of α^2. It may be the case that $S^{\langle \alpha^1, w \rangle}$ is not small enough to bound this function below by S, and using the first component to bound below by $S^{\langle \alpha^1, w\rangle}$ is not good enough to get us what we want. This means the supporting hyperplane is too close to α^1. In this case, we show that there is an interval $[b, 4b^{-1}]$ depending only on the polynomials ϕ_F over compact faces of $N(\phi)$ so that $|-2Sx^{\alpha^2} + S^{\langle \alpha^1, w\rangle} x^{\alpha^1}| = |-2Su^{\alpha^2} + Su^{\alpha^1}| \gtrsim S$, where the implicit constant is uniform, depending only on the polynomials ϕ_F. I.e., we show that we can "move"to a hyperplane containing a face of smaller codimension on which we get the desired bound. This is a recursive process, necessarily ending when the hyperplane intersects a codimension 1 face.

4.2 Supporting hyperplanes of $N(\phi)$ and scaling

The following proposition is used to define some constants necessary for applying nondegeneracy to (4.1.1). It tells us that we can move from one hyperplane not containing all vectors from some set to a new hyperplane that does contain them and such that some scaling holds with respect to the new hyperplane. Below, one should think of $v^1, \ldots, v^n$ as vertices of a compact face of a Newton polyhedron with $\langle v^i, w\rangle$ being very close or equal to 1 in the sense that $C \leq S^{\langle v^i, w\rangle - 1} \leq 1$; here we are trying to mathematically express what the example above was vaguely saying about hyperplanes being very close to some vertex. For instance, maybe the hyperplane contains $v^1, \ldots, v^{n-1}$ but not v^n. In this case we can move to a hyperplane that contains all n vertices, at the cost of estimating over a bigger box $[b, 4b^{-1}]^d$.

Proposition 4. *Let $S > 0$ and $0 < C \leq 1$. Let $v^1, \ldots, v^n \in \mathbb{R}^d$ be linearly independent* *vectors satisfying* $\langle v^i, w\rangle \geq 1$ *for all* $1 \leq i \leq n \leq d$. *Let* $\eta_i = \langle v^i, w\rangle - 1 \geq 0$ *and assume that* $C \leq S^{\eta_i} \leq 1$ *for all* i. *Let* $x \in [1,4]^d$. *There is some* $b \in (0,1)$ *depending only on* $v^1, \cdots, v^n$ *and* C, *such that there is some* $y \in [b, 4b^{-1}]^d$ *satisfying the equalities*

$$y^{v^i} = S^{\eta_i} x^{v^i}. \tag{4.2.1}$$

Furthermore, there is some positive constant ρ *depending only on* $v^1, \ldots, v^n$ *so that we can take* $b = C^{d\rho}$.

Proof. Let V be the $n \times d$ matrix with rows $v^1, \ldots, v^n$. Let $\sigma_i \in \mathbb{R}$ for $1 \leq i \leq n$ be indeterminate. Without loss of generality, assume that the first n columns of V are linearly independent and define the $d-$tuples $\sigma = (\sigma_1, \ldots, \sigma_n, 0, \ldots, 0)$ and $\eta = (\eta_1, \ldots, \eta_n, 0, \ldots, 0)$. Solving the equation $V\sigma = \eta$ can be reduced to solving $\tilde{V}\tilde{\sigma} = \tilde{\eta}$ where $\tilde{\sigma} = (\sigma_1, \ldots, \sigma_n)$, $\tilde{\eta} = (\eta_1, \ldots, \eta_n)$ and $\tilde{V} = \{v^i_j\}_{1\leq i,j\leq n}$. Since $\tilde{V}$ has full rank, we can solve $\tilde{\sigma} = \tilde{V}^{-1}\tilde{\eta}$. Denoting $\|\tilde{V}^{-1}\|_\infty = \rho$, we bound

$$\|\tilde{\sigma}\|_\infty \leq \|\tilde{V}^{-1}\|_\infty \|\tilde{\eta}\|_1 = \rho\|\tilde{\eta}\|_1.$$

Since η_i are nonnegative,

$$-\rho(\eta_1 + \cdots + \eta_n) \leq \sigma_i \leq \rho(\eta_1 + \cdots + \eta_n).$$

We can use these bounds to estimate each $S^{\sigma_i} x_i$ and find precisely which bigger box we are looking for. We use the inequalities $C \leq S^{\eta_i} \leq 1$ to bound

$$C^{d\rho} \leq (S^{\eta_1} \cdots S^{\eta_n})^{\rho} \leq 1 \leq (S^{\eta_1} \cdots S^{\eta_n})^{-\rho} \leq C^{-d\rho}.$$

Therefore

$$C^{d\rho} \leq S^{\sigma_i} x_i \leq 4C^{-d\rho}.$$

Hence, letting $b = C^{d\rho} \in (0,1)$, we see that $y \in [b, 4b^{-1}]^d$ defined by

$$y = S^{\sigma} x$$

satisfies the system of equations (4.2.1) because

$$y^{v^i} = (S^{\sigma} x)^{v^i} = S^{\langle v^i, \sigma \rangle} x^{v^i} = S^{\eta_i} x^{v^i}.$$

$\square$

Assume the hyperplane under consideration is close to, but not containing, some v^n. The results below show there do not exist other points in $\mathrm{supp}(\phi)$ on the same face as v^n far away from the hyperplane.

Proposition 5. *Let $x \in [1,4]^d$, let $\eta_1, \ldots, \eta_n \in \mathbb{R}$, and let $S > 0$. Assume η is the linear combination $\eta = \sum_{i=1}^{n} \lambda_i \eta_i$. If $v^i \in \mathbb{R}^d$ satisfy $y^{v^i} = S^{\eta_i} x^{v^i}$ for all $1 \leq i \leq n$, then $y^v = S^{\eta} x^v$ where $v = \sum_{i=1}^{n} \lambda_i v^i$.*

Proof. This is simply because

$$y^v = \prod_{i=1}^{n} y^{\lambda_i v^i} = \prod_{i=1}^{n} \left(S^{\eta_i} x^{v^i}\right)^{\lambda_i} = \prod_{i=1}^{n} S^{\lambda_i \eta_i} x^{\lambda_i v^i} = S^{\eta} x^{v}.$$

□

In particular, the scaling holds for all v in the affine hull[16] of $v^1, \ldots, v^n$. Since the sum over F in (4.1.2) can contain affinely dependent vectors, we need (5).

4.3 The main proposition: avoiding resolution of singularities

Motivated by Proposition 4, we define constants required to talk about scaling over faces $F \subset N(\phi)$ in order to apply nondegeneracy. For the rest of the chapter, we fix m such that $\phi = P_m + R_m$ and $N(\phi) = N(P_m)$.

For any codimension 1 face F of $N(\phi)$ and linearly independent $v^1, \cdots, v^n \in \operatorname{supp}(P_m) \cap F$, define V to be the $n \times d$ matrix with rows v^i and for each V pick a full rank $n \times n$ submatrix $\tilde{V}$, defined by taking n independent columns of V. Define the constant

$$\rho = \max_{V} \|\tilde{V}^{-1}\|_{\infty} \in (0, \infty),$$

where the maximum is taken over all finitely many V ($\operatorname{supp}(P_m)$ is finite).

[16]The affine hull of $\{v^1, \ldots, v^n\}$ is the set $\{\sum_{i=1}^{n} \lambda_i v^i : \sum_{i=1}^{n} \lambda_i = 1\}$.

Define the positive real number a to be the supremum

$$a = 2 \sup_{\substack{1 \le i \le d, \\ x \in [1,4]^d}} \sum_{|\alpha| \le m} \alpha_i |c_\alpha| x^\alpha \ge 2 \sup_{x \in [1,4]^d} \|x \nabla P_m(x)\|_\infty.$$

We define the constant C_0:

$$C_0 = \min_{\substack{F \subset N(\phi) \\ \text{compact}}} \inf_{x \in [1,4]^d} \|x \nabla \phi_F(x)\|_\infty.$$

By nondegeneracy, the infimum over $[1,4]^d$ over each compact face F defines some positive constant. Since there are finitely many compact faces, C_0 exists and is nonzero. We observe that $C_0 < a$ because

$$C_0 \le \inf_{x \in [1,4]^d} \|x \nabla P_m(x)\|_\infty \le a/2.$$

We used that m is large enough so that $\phi_F = (P_m)_F$. Now for $1 \le j \le d$, recursively define the constants

$$b_j = \left(\frac{C_{j-1}}{a}\right)^{d\rho}, \tag{4.3.1}$$

$$C_j' = \min_{\substack{F \subset N(\phi) \\ \text{compact}}} \inf_{y \in [b_j, 4b_j^{-1}]^d} \|y \nabla \phi_F(y)\|_{\ell^\infty(\mathbb{R}^d)},$$

and finally,

$$C_j = \min\{C'_j, C_{j-1}/a\}. \tag{4.3.2}$$

Using the convention $b_0 = 1$, it is easy to see that $C_0 > C_1 > \cdots > C_{d-1} > C_d > 0$ and therefore $b_0 > b_1 > \cdots > b_{d-1} > b_d > 0$.

If $u \in N(\phi)$ does not lie in a compact face, then we can write

$$u = v_u + \gamma_u \tag{4.3.3}$$

for some v_u lying in a compact face and $\gamma_u \in \mathbb{R}^d$, by definition of polyhedron. Since $\mathrm{supp}(P_m)$ is finite, we can define

$$p = \min_{u \in \mathrm{supp}(P_m)} ||\gamma_u||_\infty > 0. \tag{4.3.4}$$

Note that $p \leq 1$. We need to define one last constant used in the proof of the main proposition. Let

$$\delta' = \min_{\alpha^1, \alpha^2 \in \mathrm{supp}(P_m)} \inf_w \langle \frac{\alpha^1 + \alpha^2}{2}, w \rangle - 1$$

where the minimum is taken over all α^1, α^2 not contained in the same codimension 1 face, and the infimum is taken over all normals w of $N(\phi)$. In the case where there exist such α^1, α^2, we claim that $\delta' > 0$. First recall all $\alpha \in N(\phi)$ and all normals w to $N(\phi)$ satisfy $\langle \alpha, w \rangle \geq 1$. Since $N(\phi)$ is convex, α^1 and α^2 must lie on some nontrivial line segment contained in $N(\phi)$. If $\delta' = 0$ then $\langle \alpha^1, w\rangle = \langle \alpha^2, w\rangle = 1$, so any convex combination of α^1 and α^2 satisfies $\langle \lambda_1\alpha^1 + \lambda_2\alpha^2, w\rangle = 1$. This implies α^1, α^2 lie on the same codimension 1 face. The fact that $\delta' > 0$ is intuitive because the average of α^1, α^2 not lying on the same codimension 1 face lies in the interior of $N(\phi)$, and we know all points ξ in the interior satisfy $\langle \xi, w\rangle > 1$. If there are no such α^1, α^2, use the convention $\delta' = 1$. Define

$$\delta = \min\{p, \delta' p\}. \tag{4.3.5}$$

Now we are ready to set up the main proposition required to estimate $y\nabla\phi(y)$. We use δ to keep track of how small are $\langle v, w\rangle - 1$ for $v \in \operatorname{supp}(\phi) - H_w$.

Proposition 6. *Let $P \not\equiv 0$ be a polynomial defined on $[0,4]^d$ such that $P(\mathbf{0}) = 0$. Let $x \in [1,4]^d$. In terms of P, define the constants b_i, C_i, p and δ as in (4.3.1), (4.3.2), (4.3.4), and (4.3.5) respectively, for $0 \le i \le d$. Fix a normal w of $N(P)$ and let $0 < S < (C_d/a)^{\frac{1}{\delta}}$ be such that $||S^w||_\infty < (C_d/a)^{\frac{d}{\delta}}$. Then, there is a vector $\sigma \in \mathbb{R}^d_{\geq}$, a compact face $F' \supseteq F_0$, and $0 \le j' \le d$ such that*

(i) For all $v \in F'$ we have the scaling

$$S^{\langle v,w\rangle - 1}x^v = (S^\sigma x)^v, \text{ where } S^\sigma x \in [b_{j'}, 4b_{j'}^{-1}]^d, \text{ and}$$

(ii) for all $u \in \text{supp}(P) - F'$ *we have the upper bound*

$$S^{\langle u,w\rangle - 1} \leq C_{j'}/a.$$

Proof. Let $F_0 = H_w \cap N(\phi)$. First, if every $u \notin F_0$ satisfies $S^{\langle u,w\rangle-1} \leq C_0/a$, we are done with (ii) by picking $\sigma = \mathbf{0}$ and $j' = 0$. In this case (i) is also easy to see since $\langle v, w\rangle = 1$ for all $v \in F_0$, with $\sigma = \mathbf{0}$. Otherwise, for $0 \leq j \leq d-1$ define

$$\Lambda_j = \{u \in \text{supp}(P) : S^{\langle u,w\rangle-1} > C_j/a\}.$$

Each Λ_j is nonempty because $\Lambda_j \cap F_0 \neq \varnothing$. Assume $0 \leq j < d-1$ is such that $\Lambda_j = \Lambda_{j+1}$. Let us first show that Λ_j is contained in a single codimension 1 face. If some $u^1, u^2 \in \Lambda_j$ do not lie in the same codimension 1 face, we know

$$\langle u^1 + u^2, w\rangle - 2 \geq 2\delta' \geq 2\delta/p.$$

Therefore $\langle u_i, w\rangle - 1 \geq \delta/p$ for $i = 1$ or 2. Since $S^{\delta/p} \leq S^{\delta} < C_d/a < C_j/a$, we cannot have such u^i in Λ_j; this implies Λ_j is contained in a single codimension 1 face of $N(P)$. Note that if we cannot find $u^1, u^2 \in \text{supp}(P)$ not lying in the same codimension 1 face, Λ_j vacuously lies in some codimension 1 face.

Next, we show that Λ_j is contained in some compact face F_j of $N(\phi)$. If not, there is some $u \in \Lambda_j$ not lying in any compact face. By definition of polyhedron, take

$u = v_u + \gamma_u$ as in (4.3.3). Since $v_u \in \partial N(P)$, there is $1 \le i \le d$ such that

$$S^{\langle u,w\rangle - 1} = S^{\langle v,w\rangle - 1 + \langle \gamma, w\rangle} \le S^{\langle v,w\rangle - 1 + pw_i} \le S^{pw_i} \le (C_d/a)^{pd/\delta} \le C_d/a < C_{j+1}/a.$$

Assume $F_j \supset F_0$ containing Λ_j has maximal codimension, i.e., there is an affine basis $\{v^1, \ldots, v^{\dim(F_j)+1}\} \subset F_j \cap \Lambda_j$ for the affine hull of Λ_j. There is no loss of generality because a compact face cannot contain an unbounded face. By Proposition 4, we know there is some $d-$tuple σ^j so that for all $1 \le i \le \dim(F_j) + 1$ we have the equalities

$$S^{\langle v^i,w\rangle - 1} x^{v^i} = (S^{\sigma^j} x)^{v^i},$$

where $S^{\sigma^j} x \in [b_{j+1}, 4b_{j+1}^{-1}]^d$. By the definition of b_{j+1}, we can apply Proposition 4, since $S^{\langle v^i,w\rangle - 1} > C_j/a$. Proposition 5 tells us that for *all* $v \in F_j$ we have

$$S^{\langle v,w\rangle - 1} x^v = (S^{\sigma^j} x)^v.$$

Since this holds over all $0 \le j \le d-1$, we are left with claim (ii). This claim is obvious, letting $j' = j + 1$ since we assumed $\Lambda_j = \Lambda_{j+1}$, and applying Propositions 4 and 5. If there is no $0 < j < d - 1$ such that $\Lambda_j = \Lambda_{j+1}$, notice that $\dim(F_0), \dim(F_1), ...$ is a strictly increasing list of natural numbers bounded strictly above by d, and in particular, $\dim(F_j) \ge j$ for all $j \ge 0$. In this case we see that $\dim(F_{d-1}) = d - 1$. Therefore $j' = d$ satisfies property (i); property (ii) is obvious by definition of δ' and the bounds assumed on S^w. This completes the proof. $\square$

With Proposition 6 we can finish the proof of Lemma 1. As alluded to at the beginning of the chapter by proving Proposition 3, the scalings we consider over each face $F = H_w \cap N(\phi)$ are S^w for all such w, all small S. We now return to the sums (4.1.2). For all $1 \leq i \leq d$ we can write the i^{th} component of $y\nabla\phi(y)$ for $y = S^w x$ as

$$\sum_{\alpha \in F_{j'}} \alpha_i c_\alpha (S^w x)^\alpha + \sum_{\alpha \notin F_{j'}} \alpha_i c_\alpha (S^w x)^\alpha + \sum_{|\alpha|=m} h_{\alpha,\mathbf{e}_i}(S^w x)(S^w x)^\alpha$$

$$= \sum_{\alpha \in F_{j'}} \alpha_i c_\alpha S^{\langle \alpha, w\rangle} x^\alpha + \sum_{\alpha \notin F_{j'}} \alpha_i c_\alpha S^{\langle \alpha, w\rangle} x^\alpha + \sum_{|\alpha|=m} h_{\alpha,\mathbf{e}_i}(S^w x) S^{\langle \alpha, w\rangle} x^\alpha$$

$$\overset{\text{Propositions 4, 5}}{=} S\Big(\sum_{\alpha \in F_{j'}} \alpha_i c_\alpha (S^{\sigma^{j'}} x)^\alpha$$

$$+ \sum_{\alpha \notin F_{j'}} \alpha_i c_\alpha S^{\langle \alpha, w\rangle - 1} x^\alpha + \sum_{|\alpha|=m} h_{\alpha,\mathbf{e}_i}(S^w x) S^{\langle \alpha, w\rangle - 1} x^\alpha \Big). \qquad (4.3.6)$$

For the leftmost sum in the last equality, nondegeneracy guarantees that there is some $1 \leq i \leq d$ such that

$$\left| \sum_{\alpha \in F_{j'}} \alpha_i c_\alpha (S^{\sigma^{j'}} x)^\alpha \right| \geq C'_{j'}, \qquad (4.3.7)$$

since $S^{\sigma^{j'}} x \in [b_{j'}, 4b_{j'}^{-1}]$. For the second term, Proposition 6 guarantees $S^{\langle \alpha, w\rangle - 1} \leq C_{j'}/a$ for monomials $\alpha \notin F_{j'}$ appearing in P_m. Also, by the definition of a, we know

$$\sum_{\alpha \notin F_{j'}} \alpha_i |c_\alpha| x^\alpha \leq a/2. \qquad (4.3.8)$$

To estimate the remainder term, let $\zeta > 0$ be small enough so that for all $||\varepsilon||_{\ell^\infty} < \zeta$ and all $|\gamma| \leq m$ we can bound

$$\sum_{|\alpha|=m-|\gamma|} |h_{\alpha,\gamma}(4\varepsilon)| \leq \frac{C_d}{4^{m+|\gamma|}}. \tag{4.3.9}$$

This is possible since each $h_{\alpha,\gamma}$ and all summands of derivatives of R_m up to order m go to 0 as $y \to \mathbf{0}$. In particular, ζ is such that $||\varepsilon||_{\ell^\infty} < \zeta$ implies

$$\sum_{|\alpha|=m-1} |h_{\alpha,\mathbf{e}_i}(4\varepsilon)| \leq \frac{C_d}{4^{m+1}}.$$

Assuming S^w is small enough and applying the triangle inequality, along with the bounds (4.3.7), (4.3.8), and (4.3.9), we want to bound the absolute value of the sums inside the parentheses of (4.3.6), namely the quantity

$$\left| \sum_{\alpha \in F_{j'}} \alpha_i c_\alpha (S^{\sigma^{j'}} x)^\alpha + \sum_{\alpha \notin F_{j'}} \alpha_i c_\alpha S^{\langle \alpha, w \rangle - 1} x^\alpha + \sum_{|\alpha|=m} h_{\alpha,\mathbf{e}_i}(S^w x) S^{\langle \alpha, w \rangle - 1} x^\alpha \right|. \tag{4.3.10}$$

Now when we apply the triangle inequality, we need to use the fact that all monominials x^α appearing in the remainder are such that $\alpha \in N(\phi)$. We need this assumption to guarantee α does not lie below any supporting hyperplane of $N(\phi)$, in particular, $\langle \alpha, w \rangle \geq 1$, so that $S^{\langle \alpha, w \rangle - 1} \leq 1$. Here is the only piece of this proof we use either ϕ is analytic or $N(\phi)$ intersects each coordinate axis. Applying the (reverse) triangle inequality to (4.3.10), we bound it below by

$$\begin{aligned}
&\geq \left| \sum_{\alpha \in F_{j'}} \alpha_i c_\alpha (S^{\sigma^{j'}} x)^\alpha \right| - C_{j'}/a \sum_{\alpha \notin F_{j'}} \alpha_i |c_\alpha| x^\alpha - \sum_{|\alpha|=m} |h_{\alpha,\mathbf{e}_i}(S^w x) x^\alpha| \\
&\geq C'_{j'} - \frac{C_{j'}}{a} \cdot \frac{a}{2} - 4^m \frac{C_d}{4^{m+1}} \\
&= C'_{j'} - \frac{C_{j'}}{2} - \frac{C_d}{4}.
\end{aligned}$$

By definition, $C_d < C_{j'} \leq C'_{j'}$, therefore (4.3.6) is bounded below by

$$S \frac{C_d}{4} \gtrsim S. \tag{4.3.11}$$

To complete the proof of Lemma 1, let $s = \min\{(C_d/a)^{d/\delta}, \zeta\}$ where ζ was defined in (4.3.9) to ensure $h_{\alpha,\mathbf{e}_i}$ were small enough. Let $\varepsilon \in (\mathbf{0}, \mathbf{s})$. Let $y \in [\varepsilon, 4\varepsilon]$. For some $x \in [1,4]^d$, we can write $y = \varepsilon x$. By Proposition 3, there are unique $S \in (0, s^{1/d})$ and w normal to $N(\phi)$ such that $\varepsilon = S^w$. Applying the lower bound (4.3.11), we get

$$||y \nabla \phi(y)|| \gtrsim S = \varepsilon^\alpha$$

for all $\alpha \in H_w \cap N(\phi)$. Since all $\beta \in N(\phi)$ satisfy $\langle \beta, w \rangle \geq 1 = \langle \alpha, w \rangle$, we conclude $S \geq \varepsilon^\beta$:

$$S = S^{\langle \alpha, w \rangle} \geq S^{\langle \beta, w \rangle} = (S^w)^\beta = \varepsilon^\beta.$$

This finishes the proof of Lemma 1, summarizing again for future use that ε small enough means

$$\varepsilon \in (\mathbf{0}, \mathbf{s}) = (0, \min\{(C_d/a)^{d/\delta}, \zeta\})^d. \tag{4.3.12}$$

Chapter 5

Proof of Lemma 2

5.1 Estimating an integration by parts operator

Let $\phi \in C^k$ be $m-$convenient or analytic on $[-4,4]$ for some $m \geq 1$ and assume ϕ is nondegenerate. Let $\eta \in C^k$ be supported in $[1,4]^d$. The goal of this chapter is to integrate

$$I_+(\lambda, \varepsilon) = \int_{[1,4]^d} e^{i\lambda\phi(\varepsilon x)}\eta(x)dx$$

by parts k times (where we integrate by parts any number of times if ϕ is smooth) in order to get good estimates on $I_+(\lambda, \varepsilon)$ for ε small enough .

Let $f(x) = \frac{\nabla\phi(x)}{\|\nabla\phi(x)\|^2}$ for $x \in [1,4]$. Since $\nabla\phi(x) \neq 0$ away from coordinate axes, $f(x)$ is C^{k-1} in each component. We define the operator $D = D_{\varepsilon,\phi}$ on C^1 functions

$g : \mathbb{R}^d \to \mathbb{R}$ by

$$D(g)(x) = \frac{\nabla g(x) \cdot f(\varepsilon x)}{i\lambda}. \tag{5.1.1}$$

We can check that $e^{i\lambda\phi(\varepsilon x)}$ is an eigenfunction of D, which is one of the main reasons we consider this operator. If g is C^k, we can estimate $(D^t)^N(g)(x)$ for $1 \leq N \leq k$, where the adjoint D^t of D is given by the divergence

$$D^t(g)(x) = -\nabla \cdot \frac{g(x)f(\varepsilon x)}{i\lambda}. \tag{5.1.2}$$

To estimate $(D^t)^N(g)$ we consider the components f_n of f. The goal is to show $\partial^\beta f_n(x)$ is a linear combination of terms of the form

$$\frac{\partial^{\gamma^1}\phi(x) \cdots \partial^{\gamma^{2^r-1}}\phi(x)}{\|\nabla\phi(x)\|^{2^r}} = \frac{\partial^{\gamma^1}\phi(x)}{\|\nabla\phi(x)\|} \cdots \frac{\partial^{\gamma^{2^r-1}}\phi(x)}{\|\nabla\phi(x)\|} \cdot \|\nabla\phi(x)\|^{-1}, \tag{5.1.3}$$

where $0 \leq |\gamma^\ell| \leq N$. If this holds, we could bound

$$|f_n(\varepsilon x)| \lesssim \varepsilon^{-\alpha} \tag{5.1.4}$$

for any $\alpha \in N(\phi)$, for ε small enough: (5.1.3) implies $\partial^\beta f_n(\varepsilon x)$ is a linear combination of products of

$$\varepsilon^{\gamma^\ell}\partial^{\gamma^\ell}\phi(\varepsilon x)\|\varepsilon\nabla\phi(\varepsilon x)\|^{-1} \tag{5.1.5}$$

for $1 \leq \ell \leq 2^r - 1$, times $\|\varepsilon\nabla\phi(\varepsilon x)\|^{-1}$. We claim the first $2^r - 1$ terms can be bounded above by a constant independent of ε, while the last term, $\|\varepsilon\nabla\phi(\varepsilon x)\|^{-1}$, we know is bounded above by $\varepsilon^{-\alpha}$ for any $\alpha \in N(\phi)$ by Lemma 1. The claim is easy to verify by **P2**: for any ε small enough, $\varepsilon^{\gamma^\ell}\partial^{\gamma^\ell}\phi(\varepsilon x) = x^{-\gamma^\ell}(\varepsilon x)^{\gamma^\ell}\partial^{\gamma^\ell}\phi(\varepsilon x)$ is bounded above in absolute value by a uniform constant times ε^v for *some* $v \in N(\phi)$ and $x \in [1,4]^d$. Lemma 1 then guarantees the first $2^r - 1$ terms of (5.1.3) evaluated at εx, namely the terms (5.1.5), are indeed bounded above by a constant independent of ε since $\|\varepsilon\nabla\phi(\varepsilon x)\|^{-1} \lesssim \|\varepsilon x\nabla\phi(\varepsilon x)\|^{-1} \lesssim \varepsilon^{-\alpha}$ for *all* $\alpha \in N(\phi)$, so in particular for $\alpha = v$.

We proceed by examining some derivatives necessary to prove (5.1.3). Consider $\|\nabla\phi(x)\|^{2^r}$, a sum of products of $2 \cdot 2^{r-1} = 2^r$ functions, each of which is equal to some derivative of ϕ of order no more than 1. Its partial derivative with respect to x_j is

$$\partial^{\mathbf{e}_j}\|\nabla\phi(x)\|^{2^r} = 2^r\|\nabla\phi(x)\|^{2^r-2}\sum_{\ell=1}^{d}\phi'_{x_\ell}(x)\phi''_{x_\ell x_j}(x),$$

which is a sum of products of $(2^r - 2) + 2 = 2^r$ functions, each equal to some partial derivative of ϕ of order no more than 2, more precisely, $2(2^{r-1} - 1)$ such functions from the norm on the left and 2 more from the chain rule – the sum on the right.

Writing $\beta = \gamma^1 + \cdots + \gamma^{2^r-1}$, the function

$$\partial^{\mathbf{e}_j} \sum_{\beta} a_\beta \partial^{\gamma^1}\phi(x)\cdots\partial^{\gamma^{2^r-1}}\phi(x)$$

$$= \sum_{\beta}\sum_{l=1}^{2^r-1} a_\beta \partial^{\gamma^1}\phi(x)\cdots\partial^{\gamma^l+\mathbf{e}_j}\phi(x)\cdots\partial^{\gamma^{2^r-1}}\phi(x)$$

is again a sum of products of $2^r - 1$ functions, each equal to some partial derivative of ϕ of order at most one more than $|\beta|$. Therefore the numerator of

$$\partial^{\mathbf{e}_j} \frac{\sum_\beta \alpha_\beta \partial^{\gamma^1}\phi(x)\cdots\partial^{\gamma^{2^r-1}}\phi(x)}{\|\nabla\phi(x)\|^{2^r}}$$

is equal to

$$\|\nabla\phi(x)\|^{2^r} \sum_{\beta}\sum_{l=1}^{2^r-1} a_\beta \partial^{\gamma^1}\phi(x)\cdots\partial^{\gamma^l+\mathbf{e}_j}\phi(x)\cdots\partial^{\gamma^{2^r-1}}\phi(x)$$

$$-\sum_{\beta} a_\beta \partial^{\gamma^1}\phi(x)\cdots\partial^{\gamma^{2^r-1}}\phi(x)\cdot 2^r\|\nabla\phi(x)\|^{2^r-2}\sum_{\ell=1}^{d}\phi'_{x_\ell}(x)\phi''_{x_\ell x_j}(x).$$

After reorganizing, we see that we get a sum of products of $2^r + 2^r - 1 = 2^{r+1} - 1$ functions, each equal to some partial derivative of ϕ. We are left with the denominator of the partial derivative in the j direction: $\|\nabla\phi(x)\|^{2^{r+1}}$. So by induction, the proof of (5.1.4) is complete: we let $|\beta| = r > 0$ above and write $\beta = \beta' + \mathbf{e}_j$ for any j such that $\beta_j \neq 0$ (the base case $\beta = \mathbf{0}$ holds trivially).

We can now compute by induction without much work that for $\beta^0, \ldots, \beta^N \in \mathbb{R}^d$

there are constants $a_\beta = a_{\beta^0,\dots,\beta^N} \in \{0,1\}$ such that

$$(D^t)^N(g)(x) = (i\lambda)^{-N} \sum_{\substack{1\le j_1,\dots,j_N\le d\\ |\beta^0+\beta^1+\cdots+\beta^N|=N}} a_\beta \partial^{\beta^0} g(x)(\partial^{\beta^1} f_{j_1})(\varepsilon x)\cdots(\partial^{\beta^N} f_{j_N})(\varepsilon x).$$

By (5.1.4),

$$|(D^t)^N(g)(x)| \le \lambda^{-N} \sum_{\substack{1\le j_1,\dots,j_N\le d\\ |\beta^0+\beta^1+\cdots+\beta^N|=N}} a_\beta |\partial^{\beta^0} g(x)|\cdot|(\partial^{\beta^1} f_{i_1})(\varepsilon x)|\cdots|(\partial^{\beta^N} f_{i_N})(\varepsilon x)|$$

$$\lesssim \lambda^{-N} \sum_{j_1,\dots,j_n=1}^{d} |\partial^{\beta^0} g(x)|\varepsilon^{-\alpha^1}\cdots\varepsilon^{-\alpha^N} \tag{5.1.6}$$

for any $\alpha^1,\cdots,\alpha^N \in N(\phi)$. In particular, for all $\alpha \in N(\phi)$,

$$|(D^t)^N(g)(x)| \lesssim \lambda^{-N} \max_{1\le|\beta^0|\le N} |\partial^{\beta^0} g(x)|\varepsilon^{-N\alpha} \tag{5.1.7}$$

for all $1 \le N \le k$, where the implicit constant is independent of ε and λ. Note that we also trivially have an estimate for $N = 0$. All that is left to prove Lemma 2 is to integrate by parts.

5.2 Final estimate for Lemma 2

We now put everything together for ε small enough ($||\varepsilon||_\infty \le \mathbf{s}$ from (4.3.12)):

$$I_+(\lambda,\varepsilon) = \int_{[\varepsilon,4\varepsilon]} e^{i\lambda\phi(x)} x^\beta \eta_\varepsilon(x)dx = \varepsilon^{\mathbf{1}} \int_{[1,4]^d} e^{i\lambda\phi(\varepsilon x)}(\varepsilon x)^\beta \eta(x)dx$$

$$= \varepsilon^{\beta+\mathbf{1}} \int_{[1,4]^d} D^N(e^{i\lambda\phi(\varepsilon x)})x^\beta \eta(x)dx = \varepsilon^{\beta+\mathbf{1}} \int_{[1,4]^d} e^{i\lambda\phi(\varepsilon x)}(D^t)^N(\cdot^\beta \eta)(x)dx.$$

By (5.1.7), letting $g(x) = x^\beta \eta(x) \in C^k$ in (5.1.7),

$$\int_{[1,4]^d} |(D^t)^N(\cdot^\beta \eta)(x)|dx \lesssim \int_{[1,4]^d} \lambda^{-N}\varepsilon^{-N\alpha}dx \lesssim \lambda^{-N}\varepsilon^{-N\alpha}.$$

Therefore we have proved Lemma 2: for all $1 \le N \le k$,

$$|I_+(\lambda,\varepsilon)| \lesssim \lambda^{-N}\varepsilon^{-(N\alpha-\beta-\mathbf{1})}.$$

5.3 Application of (5.1.7) to other amplitudes

In the proof of Theorem 2, we need to estimate the integrals

$$\int e^{i\lambda\phi(x)}x^\gamma \partial^\gamma R_m(x)x^\beta \psi(x)dx,$$

where ψ is C^k and supported close enough to the origin. In order to estimate this integral, we first estimate

$$\int_{[\varepsilon,4\varepsilon]} e^{i\lambda\phi(x)}\partial^\gamma R_m(x)x^\beta \eta_\varepsilon(x)dx = \varepsilon^{\beta+\mathbf{1}} \int_{[1,4]^d} e^{i\lambda\phi(\varepsilon x)}\partial^\gamma R_m(\varepsilon x)x^\beta \eta(x)dx$$

for some C^k compactly supported $\eta : [1,4]^d \to \mathbb{R}$. We apply the same argument as in the previous section, and apply property **P1** for remainders of analytic or

$m-$convenient C^k functions. Write $\partial^\gamma R_m(\varepsilon x) = R_{m,\gamma,\varepsilon}$ and $x^\beta \eta(x) = \eta'(x)$. By 5.1.7,

$$(D^t)^N(R_{m,\gamma,\varepsilon}\eta') \lesssim \lambda^{-N} \max_{1\leq|\beta^0|\leq N} |\partial^{\beta^0}(R_{m,\gamma,\varepsilon}\eta')|\varepsilon^{-N\alpha}$$

for all $1 \leq N \leq k - |\gamma|$. By the Leibniz formula,

$$\partial^\beta(R_{m,\gamma,\varepsilon}\eta') = \sum_{\alpha_i\leq\beta_i} \partial^\alpha R_{m,\gamma,\varepsilon}\partial^{\beta-\alpha}\eta' = \sum_{\alpha_i\leq\beta_i} \varepsilon^\alpha \partial^{\alpha+\gamma} R_m(\varepsilon x)\partial^{\beta-\alpha}\eta'$$

$$= \varepsilon^{-\gamma} \sum_{\alpha_i\leq\beta_i} \varepsilon^{\alpha+\gamma}\partial^{\alpha+\gamma} R_m(\varepsilon x)\partial^{\beta-\alpha}\eta'.$$

$$\overset{(3.2.4)}{=} \varepsilon^{-\gamma} \sum_{|v|=m}\sum_{\alpha_i\leq\beta_i} \tilde{h}_{\alpha,\gamma}(\varepsilon x)\varepsilon^v x^v \partial^{\beta-\alpha}\eta'.$$

Since $x \in [1,4]^d$, for some uniform constant independent of ε we can bound

$$|\varepsilon^{\alpha+\gamma}\partial^{\alpha+\gamma} R_m(\varepsilon x)| \lesssim \sum_{|v|=m} \varepsilon^{v-\gamma}.$$

Finally,

$$\left| \int_{[\varepsilon,4\varepsilon]^d} e^{i\lambda\phi(x)} x^\gamma \partial^\gamma R_m(x) x^\beta \eta_\varepsilon(x)dx \right| \lesssim \sum_{|v|=m} \lambda^{-N}\varepsilon^{v+\beta+\mathbf{1}-N\alpha} \tag{5.3.1}$$

for all $\alpha \in N(\phi)$, all ε small enough, and all $1 \leq N \leq k - |\gamma|$. We will go back to this formula in the next chapter to prove a corollary required for Chapter 7.

Chapter 6

Proof of Theorem 1

We now use Lemma 2 and linear programming to prove Varchenko's upper bounds. We can sum over positive j_i to get a bound on the integral

$$I_+(\lambda) = \int_{\mathbb{R}^d_\geq} e^{i\lambda\phi(x)}\psi(x)dx = \int_{[0,4]^d} e^{i\lambda\phi(x)}\psi(x)dx$$

where ψ is supported in a sufficiently small neighborhood of the origin. This is achieved by decomposing $\psi(x) = \sum_{j_1,\ldots,j_d=0}^{\infty} \psi(x)f_j(x)$, where $f_j(x)$ is a partition of unity subordinate to the cover $\{(2^{-j}, 2^{-j+\mathbf{2}})\}_{j\in\mathbb{N}^d}$ of $(0,4)^d$. One should choose a family $\{f_j\}$ for which there exists a uniform constant $C > 0$ such that

$$\left|\frac{\partial^\alpha f_j(x)}{\partial x^\alpha}\right| \leq Cx^{-\alpha},$$

so that when one scales the support of $(\psi f_j)(x)$ to $[1,4]^d$ by 2^{-j}, all derivatives of $(\psi f_j)(2^{-j}x)$ are bounded above by a uniform constant independent of j. Since it is sufficient to find some $f_0(x)$ such that the functions $f_j(x) = f_0(2^j x)$ define our partition of unity, it is not difficult to prove that indeed we can choose a family $\{f_j\}_{j\in\mathbb{N}^d}$ as required. Existence of such a partition of unity is well-known, especially for those who use Littlewood-Paley decomposition. See, for example, [3].[17]

6.1 Easy case: when $\beta = 0$ and $\mathbf{t} \in \mathbb{R}^d$ lies in a single codimension 1 face

Varchenko showed[18] that if $\phi : \mathbb{R}^d \to \mathbb{R}$ is real analytic and nondegenerate in a neighborhood of the origin, and if $\psi : \mathbb{R}^d \to \mathbb{R}$ is smooth and supported close enough to the origin, then

$$|I(\lambda)| \lesssim \lambda^{-1/t} \log^{d_j - 1}(\lambda)$$

where t is the Newton distance of ϕ and d_j is the highest codimension over all faces containing $\mathbf{t}$. Moreover, he showed this result is sharp for $t > 1$. We apply Lemma 2 to obtain a more general result at the end of the next section, although we do not prove the result is sharp. For the rest of this section, we use i as an index. Fix $\lambda > 2$. Let t be the Newton distance of ϕ and assume for simplicity that $\mathbf{t}$ lies in a

[17] In order to apply the result, one would define the function ψ generating the dyadic partition of unity for one variable, then multiply d many $\psi_1 \cdots \psi_d$ to obtain a partition of unity for dyadic cubes.

codimension 1 face $F \subset N(\phi)$ and does not lie in any face of higher codimension. This case illustrates the main ideas. By Lemma 2, it is enough to show

$$\sum_{j_1,\ldots,j_d=0}^{\infty} \min_{N\geq 0,\ \alpha\in N(\phi)} \{\lambda^{-N} 2^{\langle N\alpha-\mathbf{1},j\rangle}\} \lesssim \lambda^{-1/t}.$$

First, setting $N = 0$, we see that

$$\sum_{\substack{j_1,\ldots,\hat{j}_i,\ldots,j_d=0 \\ j_i=\log(\lambda)/t}}^{\infty} 2^{\langle -\mathbf{1},j\rangle} \lesssim \lambda^{-1/t}.$$

Hence, it is enough to bound

$$\sum_{j_1,\ldots,j_d=0}^{\log(\lambda)/t} \min_{N\geq 0,\ \alpha\in N(\phi)} \{\lambda^{-N} 2^{\langle N\alpha-\mathbf{1},j\rangle}\} \tag{6.1.1}$$

above by a uniform positive constant times $\lambda^{-1/t}$. Since $\mathbf{t}$ lies in a face of codimension 1 that cannot lie in a coordinate hyperplane ($t > 0$) there are linearly independent $\alpha^1, \ldots, \alpha^d \in F$ whose convex hull contains $\mathbf{t}$, so we write

$$\mathbf{t} = \sum_{i=1}^{d} \lambda_i \alpha^i.$$

For the rest of the proof we fix $N > 1/t$.[18] For $1 \leq i \leq d$ let $\theta_i = \frac{\lambda_i}{Nt}$ and

[18] We can assume $N > 1/t$ only if $n > 1/t$. In particular, there is no concern if ϕ is analytic or smooth convenient.

$\theta_0 = 1 - \frac{1}{Nt}$. Then all θ_i are positive and sum to 1. Moreover, we can check

$$\theta_0(-\mathbf{1}) + \sum_{i=1}^{d} \theta_i (N\alpha^i - \mathbf{1}) = \mathbf{0}.$$

We estimate (6.1.1) by considering the sum

$$\sum_{j_1,\ldots,j_d=0}^{\log(\lambda)/t} \min_{1\le i\le d} \{J_0 2^{\langle -\mathbf{1}, j\rangle}, J_i 2^{\langle N\alpha^i - \mathbf{1}, j\rangle}\}, \tag{6.1.2}$$

where $J_0 = 1$ and the coefficients J_i equal λ^{-N} for $1 \le i \le d$; (6.1.2) clearly bounds (6.1.1) above, since we fixed N. Letting A be the full rank matrix $\{\alpha^i_\ell\}_{1\le i,\ell\le d}$, we can solve $Az = \mathbf{1} \in \mathbb{R}^d$. Write the solution as $z = (z_1, \ldots, z_d)$ (we already know $z = w(\mathbf{t})$, but z is not as obvious when F is not codimension 1). Since the convex hull of $\{\alpha^1, \ldots, \alpha^d\}$ contains $\mathbf{t} \in \mathbb{R}^d$, we conclude $\langle \mathbf{t}, z\rangle = 1$, hence $\langle \mathbf{1}, z\rangle = 1/t$. Denoting the $d-$tuple $\log(\lambda)z$ by j^0, we compute

$$J_0 2^{\langle -\mathbf{1}, j^0\rangle} = \lambda^{-1/t} = J_i 2^{\langle N\alpha^i - \mathbf{1}, j^0\rangle}.$$

Hence, by reindexing and factoring out $\lambda^{-1/t}$, we see

$$\sum_{j_1,\ldots,j_d=0}^{\log(\lambda)/t} \min_{1\le i\le d} \{J_0 2^{\langle -\mathbf{1}, j\rangle}, J_i 2^{\langle N\alpha^i - \mathbf{1}, j\rangle}\}$$

$$\lesssim \lambda^{-1/t} \sum_{j_1=-\log(\lambda)z_1}^{\log(\lambda)/t-\log(\lambda)z_1} \cdots \sum_{j_d=-\log(\lambda)z_d}^{\log(\lambda)/t-\log(\lambda)z_d} \min_{1\le i\le d}\{2^{\langle -\mathbf{1},j\rangle}, 2^{\langle N\alpha^i-\mathbf{1},j\rangle}\}. \tag{6.1.3}$$

Now notice that the vectors $N\alpha^i - \mathbf{1}$ and $-\mathbf{1}$ do not all lie in the same hyperplane: $\langle -\mathbf{1}, z\rangle = -1/t$ and $\langle N\alpha_i, z\rangle = N \neq 0$. This finishes the claim, since $\{\xi \in \mathbb{R}^d : \langle \xi, z\rangle = 1\}$ is the unique hyperplane in $\mathbb{R}^d$ containing the d many linearly independent vectors α^i. Therefore, for all $x \in \mathbb{R}^d$,

$$\sup_{\|x\|_\infty=1} \min_{1\le i\le d}\{\langle -\mathbf{1}, x\rangle, \langle N\alpha^i - \mathbf{1}, x\rangle\} < 0.$$

By homogeneity, there is some $c > 0$ such that

$$\min_{1\le i\le d}\{\langle -\mathbf{1}, x\rangle, \langle N\alpha^i - \mathbf{1}, x\rangle\} \le -c\|x\|_\infty.$$

Apply this fact to bound the sum in (6.1.3) by

$$\sum_{j_1,\ldots,j_d\in\mathbb{Z}} 2^{\min\{\langle -\mathbf{1},j\rangle,\ \langle N\alpha^i-\mathbf{1},j\rangle\}} \lesssim \sum_{n=0}^{\infty} \sum_{\|j\|_\infty=n} n^{d-1}2^{-cn} \lesssim 1$$

to see that (6.1.1) is bounded above by a uniform constant times $\lambda^{-1/t}$, which is exactly the estimate we were looking for.

6.2 General case

Let F be some codimension 1 face intersecting the line $L = \{s(\beta + \mathbf{1}) : s \in \mathbb{R}\}$ and let w be the normal of F. Since $\beta + \mathbf{1}$ has all positive components, the quantity $\langle \beta + \mathbf{1}, w\rangle$ is nonzero. Since $\beta + \mathbf{1}$ lies on a line intersecting F, we conclude that the vector $v = \frac{\beta+\mathbf{1}}{\langle \beta+\mathbf{1}, w\rangle}$ lies on F. Let $r - 1$ be the lowest dimension over all faces intersected by L. By Lemma 2, it is enough to show

$$\sum_{j_1,\ldots,j_d=0}^{\infty} \min_{N\geq 0,\ \alpha\in N(\phi)} \{\lambda^{-N} 2^{\langle N\alpha-\beta-\mathbf{1}, j\rangle}\} \lesssim \lambda^{-\langle \beta+\mathbf{1}, w\rangle} \log^{d-r}(\lambda).$$

First, setting $N = 0$, we see for any j_i that

$$\sum_{j_1=0}^{\infty} \cdots \sum_{j_i=\log(\lambda)/v_i}^{\infty} \cdots \sum_{j_d=0}^{\infty} 2^{\langle -\beta-\mathbf{1}, j\rangle} \lesssim \lambda^{-(\beta_i+1)/v_i} = \lambda^{-\langle \beta+\mathbf{1}, w\rangle}.$$

Hence, it is enough to bound

$$\sum_{j_1=0}^{\log(\lambda)/v_1} \cdots \sum_{j_d=0}^{\log(\lambda)/v_d} \min_{N\geq 0,\ \alpha\in N(\phi)} \{\lambda^{-N} 2^{\langle N\alpha-\beta-\mathbf{1}, j\rangle}\} \tag{6.2.1}$$

above by a uniform positive constant times $\lambda^{-\langle \beta+\mathbf{1}, w\rangle} \log^{d-r}(\lambda)$. It is more natural to work in a continuous setting when $r \neq d$ because we need to change variables, so we bound (6.2.1) above by a uniform constant times

$$\int_0^{\log(\lambda)/v_1} \cdots \int_0^{\log(\lambda)/v_d} \min_{N\geq 0, \alpha\in N(\phi)} \{\lambda^{-N} e^{\langle N\alpha-\beta-\mathbf{1}, x\rangle}\} dx, \tag{6.2.2}$$

and we estimate the integral (6.2.2) instead. Since v lies in a face of dimension no less than $r-1$ that also does not lie in a coordinate hyperplane, there are linearly independent $\alpha^1, \ldots, \alpha^r \in F$ whose convex hull contains v, so write

$$v = \sum_{i=1}^{r} \lambda_i \alpha^i. \tag{6.2.3}$$

For the rest of the proof we assume $N > \langle \beta + \mathbf{1}, w \rangle$. This is where we require that $k > \langle \beta + \mathbf{1}, w \rangle$. For $1 \leq i \leq r$ let $\theta_i = \lambda_i \langle \beta + \mathbf{1}, w \rangle / N$ and $\theta_0 = 1 - \langle \beta + \mathbf{1}, w \rangle / N$. All θ_i are positive and their sum is 1 by the restriction placed on N. Moreover, we can check

$$\theta_0(-\beta - \mathbf{1}) + \sum_{i=1}^{r} \theta_i (N\alpha^i - \beta - \mathbf{1}) = \mathbf{0}. \tag{6.2.4}$$

The integral (6.2.2) can be bounded above by

$$\int_0^{\log(\lambda)/v_1} \cdots \int_0^{\log(\lambda)/v_d} \min_{1 \leq i \leq k} \{ e^{\langle -\beta - 1, x \rangle}, \lambda^{-N} e^{\langle N\alpha^i - \beta - \mathbf{1}, x \rangle} \} dx. \tag{6.2.5}$$

We now change variables for convenience: define the matrix A by the equalities

$$A\alpha^i = \mathbf{e}_i \text{ for } 1 \leq i \leq r,$$

and

$$A\mathbf{e}_i = \mathbf{e}_i \text{ for } r < i \leq d.$$

Note that $\langle A^T x, \alpha^i \rangle = x_i$ for $1 \leq i \leq k$. Letting $x = A^T y$, up to a factor of the Jacobian of A^{-1}, the integral (6.2.5) equals

$$\int_{\substack{0 \leq \langle A^T y, \mathbf{e}_i \rangle \leq \log(\lambda)/v_i \\ 1 \leq i \leq d}} \min_{1 \leq j \leq k} \{ e^{\langle A(-\beta - \mathbf{1}), y \rangle}, \lambda^{-N} e^{\langle A(N\alpha^j - \beta - \mathbf{1}), y \rangle} \} dy. \tag{6.2.6}$$

We have seen that $A\alpha^i = \mathbf{e}_i$ by definition of A. Moreover, by (6.2.3), since $v = \frac{\beta + \mathbf{1}}{\langle \beta + \mathbf{1}, w \rangle}$, the vector $A(\beta + \mathbf{1})$ can be written as a linear combination of vectors $A\alpha^i = \mathbf{e}_i$ over $1 \leq i \leq r$. Therefore we can integrate over directions $r < i \leq d$

$$\int_{\substack{0 \leq \langle A^T y, \mathbf{e}_i \rangle \leq \log(\lambda)/v_i \\ k < i \leq d}} dy_{r+1} \cdots dy_d \lesssim \log^{d-r}(\lambda)$$

in order to bound (6.2.6) above by

$$\log^{d-k}(\lambda) \int_{\mathbb{R}^k} \left| \min_{1 \leq i \leq k} \{ e^{\langle A(-\beta - \mathbf{1}), y \rangle}, \lambda^{-N} e^{\langle A(N\alpha^i - \beta - \mathbf{1}), y \rangle} \} \right| dy_1 \cdots dy_r. \tag{6.2.7}$$

Since $A\alpha^i = \mathbf{e}_i$ and $\sum_{j=1}^r \lambda_j = 1$, we see

$$\langle A(\beta + \mathbf{1}), \log(\lambda)\mathbf{1} \rangle = \langle \beta + \mathbf{1}, w \rangle \log(\lambda) \sum_{j=1}^{r} \lambda_j \langle A\alpha^j, \mathbf{1} \rangle = \langle \beta + \mathbf{1}, w \rangle \log(\lambda),$$

and therefore, exponentiating, we obtain $e^{\langle A(-\beta - \mathbf{1}), \log(\lambda)\mathbf{1} \rangle} = \lambda^{-\langle \beta + \mathbf{1}, w \rangle}$. This calculation inspires the change of variables $y \to y + \log(\lambda)\mathbf{1}$. After changing variables, we can factor out $\lambda^{-\langle\beta+\mathbf{1},w\rangle}$ and bound (6.2.7) above by $\lambda^{-\langle\beta+\mathbf{1},w\rangle}\log^{d-k}(\lambda)$ times

$$\int_{\mathbb{R}^k}\Big|\min_{1\le i\le k}\{e^{\langle A(-\beta-\mathbf{1}),y\rangle},e^{\langle A(N\alpha^i-\beta-\mathbf{1}),y\rangle}\}\Big|dy_r\cdots dy_1.$$

The factors λ^{-N} disappeared after changing variables because $\lambda^{-N}e^{N\log(\lambda)}=1$. By (6.2.4),

$$\mathbf{0}=\theta_0A(-\beta-\mathbf{1})+\sum_{i=1}^{r}\theta_iA(N\alpha^i-\beta-\mathbf{1}).$$

Since $A(N\alpha^i-\beta-\mathbf{1})$ for $1\le i\le r$ and $A(-\beta-\mathbf{1})$ are linearly independent in $\mathbb{R}^r$,

$$\sup_{||y||_2=1}\min_{1\le i\le r}\{\langle A(-\beta-\mathbf{1}),y\rangle,\langle A(N\alpha^i-\beta-\mathbf{1}),y\rangle\}<0.$$

By homogeneity, there is a constant $c=c(\alpha^1,\dots,\alpha^r,\beta)>0$ such that

$$\min_{1\le i\le r}\{\langle A(-\beta-\mathbf{1}),y\rangle,\langle A(N\alpha^i-\beta-\mathbf{1}),y\rangle\}<c||y||_2.$$

After a polar change of variables, we can bound (6.2.7) by a constant independent of λ (but depending on $\alpha^1,\dots,\alpha^r$, and β) times

$$\lambda^{-\langle\beta+\mathbf{1},w\rangle}\log^{d-r}(\lambda)\int_0^\infty e^{-cr}dr\lesssim\lambda^{-\langle\beta+\mathbf{1},w\rangle}\log^{d-r}(\lambda),$$

which is exactly the bound we were looking for.

Varchenko's upper bounds as a special case

Letting $\beta = \mathbf{0}$ above, the inner product $\langle \beta + \mathbf{1}, w(\beta + \mathbf{1})\rangle = \langle \mathbf{1}, w(\mathbf{1})\rangle = 1/t$ by Observation 1. Therefore Varchenko's upper bound in [18] is a special case of Theorem 1. In fact, the upper bound hold for all convenient nondegenerate phases $\phi \in C^k$, assuming $k > 1/t$. Since $t \geq 1/d$ holds for all C^k functions, one could guarantee such an estimate for all C^{d+1} convenient phases.

6.3 An estimate for remainders

Let $\phi \in C^k$ be $m-$convenient and write $\phi = P_{m'} + R_{m'}$, where we are free to choose $m \leq m' \leq n$. With the same methods used above, we want to estimate

$$I_{R,\gamma}(\lambda) = \int_{[0,4]^d} e^{i\lambda\phi(x)} x^\gamma \partial^\gamma R_m(x) x^\beta \psi(x) dx,$$

where ψ is $C^{k-|\gamma|}$ and supported close enough to the origin. We know by (5.3.1) that for all $1 \leq N \leq k - |\gamma|$,

$$\left| \int_{[\varepsilon,4\varepsilon]^d} e^{i\lambda\phi(x)} x^\gamma \partial^\gamma R_m(x) x^\beta \eta_\varepsilon(x) dx \right| \lesssim \sum_{|v|=m} \lambda^{-N} \varepsilon^{v+\beta+\mathbf{1}-N\alpha}$$

where η is $C^{k-|\gamma|}$ and supported in $[1,4]^d$. We again use linear programming over each v, assuming that $k > |\gamma| + 1 + \langle \beta + \mathbf{1}, w(\beta + \mathbf{1})\rangle$ so that $k - |\gamma| \geq N >$

$\langle v+\beta+\mathbf{1}, w(v+\beta+\mathbf{1})\rangle$ for all $|v|=m$. The estimate is

$$|I_{R,\gamma}(\lambda)| \lesssim \sum_{|v|=m} \lambda^{-\langle v+\beta+\mathbf{1}, w(v+\beta+\mathbf{1})\rangle} \log^{d-1}(\lambda) \lesssim \lambda^{-1-\langle \beta+\mathbf{1}, w(\beta+\mathbf{1})\rangle} \log^{d-1}(\lambda). \quad (6.3.1)$$

With the same techniques, we get

Corollary 1. *Let $\gamma^1, \ldots, \gamma^n, \beta \in \mathbb{N}^d$. Assume $\phi \in C^k$ is $m-$convenient. Assume*

$$k > \max_{1\leq j\leq n} |\gamma^j| + \langle v^1+\cdots+v^n+\beta+\mathbf{1}, w(v^1+\cdots+v^n+\beta+\mathbf{1})\rangle$$

for all $|v^i|=m$, and let ψ be $C^{k-\max|\gamma^i|}$ with support close enough to the origin. Let $\mu(v^1,\ldots,v^n) = \langle v^1+\cdots+v^n+\beta+\mathbf{1}, w(v^1+\cdots+v^n+\beta+\mathbf{1})\rangle$. Then

$$\left| \int_{\mathbb{R}^d} e^{i\lambda\phi}(x^{\gamma^1}\partial^{\gamma^1}R_m)\cdots(x^{\gamma^n}\partial^{\gamma^n}R_m)x^\beta\psi \right| \lesssim \sum_{|v^1|=\cdots=|v^n|=m} \lambda^{-\mu(v^1,\ldots,v^n)} \log^{d-1}(\lambda).$$

The minimum over $k-|\gamma^j|$ is the maximum number we can differentiate each of the terms $x^\gamma\partial^\gamma R_m$. Note that if ϕ is smooth convenient or analytic, we don't need to worry about how large k has to be.

Chapter 7

Theorem 2: asymptotic expansion of $I(\lambda)$

We remind the reader[19] that if some multiple of $\beta \in \mathbb{N}^d$ lies in $N(\phi)$, we define

$$w(\beta) = \{w : \langle \beta, w \rangle \text{ is minimal}\}$$

where the minimum is taken over all finitely many normals w of codimension 1 faces F of $N(\phi)$. In the same section we also defined $\langle \gamma, w(\beta) \rangle$ for the scalar $\max_{w \in w(\beta)} \langle \gamma, w \rangle$. In this chapter, we make use of the facts proven in Proposition 2 at the end of Section 3.4:

If $\langle \beta, w \rangle = \langle \alpha + \beta, w(\alpha + \beta) \rangle$, then

[19] See Chapter 3, Section 3.3

P3 $w(\alpha+\beta) \subseteq w(\beta)$ and $\alpha_j w_j = 0$ for all $w \in w(\alpha+\beta)$,

P4 either $w(\alpha+\beta) \subsetneq w(\beta)$ or else $\alpha \neq \mathbf{0}$ implies β does not lie in any compact codimension 1 face, and

P5 If $\alpha \neq \mathbf{0}$, then $|w(\alpha+\beta)| \leq |\{j : \alpha_j \neq 0\}| \leq d-1$.

P6 If instead we assume $\alpha \in N(\phi)$ and $\langle \beta, w \rangle + 1 = \langle \alpha+\beta, w(\alpha+\beta) \rangle$, then $|w(\alpha+\beta)| \leq d$ with equality only if $\alpha = \beta$.

7.1 Derivatives of $I(\lambda)$

In this section we choose $m < m' \overset{def}{=} m+n+1 \leq k$ and express ϕ as $P_{m'} + R_{m'}$ where $P_{m'}$ is a degree m' polynomial and $R_{m'} = \sum_{|\alpha|=m'} x^\alpha h_\alpha(x)$. Denote the integral $\int_{\mathbb{R}^d} e^{i\lambda\phi(x)}\psi(x)dx$ by $I(\lambda)$. We want to prove that $I(\lambda)$ has asymptotic expansion for large λ of the form

$$I(\lambda) \sim \sum_{j=0} \sum_{r=0}^{d_j-1} a_{j,r}(\psi) \lambda^{-p_j} \log^{d_j-1-r}(\lambda) \tag{7.1.1}$$

in the sense described prior to Theorem 2, where $p_0 < p_1 < \cdots$ is the ordering of the countable set $\mathcal{C}$ of Section 3.5 (Chapter 3). We assume that $k > (d_0+\cdots+d_n)m'+d$, and d_j is the greatest codimension over all faces intersecting the lines

$$\{s(\beta+\mathbf{1}) : s \in \mathbb{R},\ \beta \in \mathbb{N}^d,\ \langle \beta+\mathbf{1}, w(\beta+\mathbf{1}) \rangle - n' = p_j \text{ some } n' \in \mathbb{N}\}.$$

Note that we only need the Newton polyhedron to know what are each p_j and d_j; no additional conditions other than being able to define $N(\phi)$ are required of ϕ.

To start the proof, we need to first rewrite $P_{m'}$ in a suggestive way and then differentiate $I_\beta(\lambda)$.

If $P(x)$ is a polynomial of degree at most m', note that

$$P(x) = \sum_{|\alpha| \leq m'} c_\alpha x^\alpha = \sum_{|\alpha| \leq m'} \sum_{j=1}^{d} \alpha_j v_j c_\alpha x^\alpha,$$

where we are free to choose any $v \in \mathbb{R}^d_{\geq}$ satisfying $\langle \alpha, v \rangle = 1$; we suppress the dependence on α for notational convenience. Let $w \in w(\beta + \mathbf{1})$ be free. Recall: geometrically, $F = H_w \cap N(\phi)$ is a codimension 1 face hit by the line $\{s(\beta+\mathbf{1}) : s \in \mathbb{R}\}$, so in particular F does not lie in a coordinate hyperplane. We can rewrite $P_{m'}$ as

$$\sum_{\alpha} \sum_{j=1}^{d} \alpha_j (v_j - w_j) c_\alpha x^\alpha + \sum_{\alpha} \sum_{j=1}^{d} \alpha_j w_j c_\alpha x^\alpha. \tag{7.1.2}$$

This expression also works for arbitrary analytic functions. Letting $v = v(\alpha) = w$ for all $\alpha \in F$, the quantity (7.1.2) simplifies to

$$\sum_{\alpha \notin F} \sum_{j=1}^{d} \alpha_j (v_j - w_j) c_\alpha x^\alpha + \sum_{\alpha} \sum_{j=1}^{d} \alpha_j w_j c_\alpha x^\alpha. \tag{7.1.3}$$

Since we are integrating over a compact set and $e^{i\lambda\phi}$ has the same smoothness as ϕ, we can write $\lambda \frac{d}{d\lambda} I_\beta(\lambda) =$

$$\int e^{i\lambda\phi(x)} i\lambda\phi(x) x^\beta \psi(x) dx = \underbrace{\int e^{i\lambda\phi} i\lambda(wx \cdot \nabla\phi) x^\beta \psi}_{I_1} + \underbrace{\int e^{i\lambda\phi} i\lambda(\phi - wx \cdot \nabla\phi) x^\beta \psi}_{I_2}.$$

We first estimate I_1. Integration by parts tells us

$$I_1 = -\int e^{i\lambda\phi(x)} \nabla \cdot (x^\beta \psi(x) wx) dx.$$

By the product rule, and the definition $\langle \beta + \mathbf{1}, w \rangle = \sum_{j=1}^d (\beta_j + 1) w_j$, we can rewrite the integral above as

$$I_1 = -\langle \beta + \mathbf{1}, w \rangle \int e^{i\lambda\phi} x^\beta \psi - \underbrace{\int e^{i\lambda\phi} x^\beta (wx \cdot \nabla\psi)}_{I_{11}}.$$

Define D_β to be the operator $\lambda \frac{d}{d\lambda} + \langle \beta + \mathbf{1}, w \rangle$. We have just shown

$$D_\beta I_\beta(\lambda) = I_2 - I_{11}.$$

We begin with the estimate of

$$I_{11} = \sum_{j=1}^d w_j \int e^{i\lambda\phi} x^{\beta + \mathbf{e}_j} \psi'_{x_j}$$

By Theorem 1,

$$|I_{11}(\lambda)| \lesssim \sum_{j=1}^{d} w_j \lambda^{-\langle \beta+\mathbf{e}_j+\mathbf{1}, w(\beta+\mathbf{e}_j+\mathbf{1})\rangle} \log^{d_j-1}(\lambda). \tag{7.1.4}$$

- If $\langle \beta+\mathbf{e}_j+\mathbf{1}, w(\beta+\mathbf{e}_j+\mathbf{1})\rangle > \langle \beta+\mathbf{1}, w\rangle$, we are done with the estimate.

- Otherwise, by property **P3** we can say $w(\beta+\mathbf{1}+\mathbf{e}_j) \subseteq w(\beta+\mathbf{1})$. If equality holds, **P4** guarantees $w_j = 0$; otherwise, we have a few more cases.

- If $s(\beta+\mathbf{1})$ doesn't intersect $N(\phi)$ at a vertex, then $|w(\beta+\mathbf{1})|$ is the highest codimension over all faces containing $\beta+\mathbf{1}$. In this case, $|w(\beta+\mathbf{1}+\mathbf{e}_j)| < |w(\beta+\mathbf{1})|$ so $\beta+\mathbf{1}+\mathbf{e}_j$ lies in a strictly smaller codimension face.

- If $s(\beta+\mathbf{1})$ intersects in a vertex and $s(\beta+\mathbf{1}+\mathbf{e}_j)$ does not, the power of log again must be smaller.

- Finally, **P5** tells us the lines $s(\beta+\mathbf{1})$ and $s(\beta+\mathbf{1}+\mathbf{e}_j)$ cannot both intersect $N(\phi)$ at vertices.

We have finally finished proving an estimate on (7.1.4) that are summarized succinctly at the end of the section.

In order to estimate I_2, we first rewrite

$$I_2 = \underbrace{\int e^{i\lambda\phi} i\lambda(P_{m'} - wx\cdot\nabla P_{m'})x^\beta\psi}_{I_{21}} + \underbrace{\int e^{i\lambda\phi} i\lambda(R_{m'} - wx\cdot\nabla R_{m'})x^\beta\psi}_{I_{22}}.$$

By (7.1.3), $P_{m'}(x) - wx \cdot \nabla P_{m'}(x) = \sum_{\alpha \notin F} \sum_{j=1}^{d} (v_j - w_j) c_\alpha x^\alpha$. If this quantity is zero, we are done with the estimate. Otherwise, Theorem 1 tells us we can bound I_{21} above by

$$|I_{21}| \lesssim \lambda \cdot \lambda^{-\langle \alpha+\beta+\mathbf{1}, w(\alpha+\beta+\mathbf{1})\rangle} \log^{d'-1}(\lambda)$$

for d' guaranteed by Theorem 1. Note that

$$\langle \beta + \mathbf{1}, w\rangle + \langle \alpha, w(\alpha + \beta + \mathbf{1})\rangle \leq \langle \alpha + \beta + \mathbf{1}, w(\alpha + \beta + \mathbf{1})\rangle.$$

Moreover, $\langle \alpha, w(\alpha + \beta + \mathbf{1})\rangle \geq 1$ since α must lie on or above $H_{w'}$ for all $w' \in w(\alpha+\beta+\mathbf{1})$ by convexity of $N(\phi)$. If $1 + \langle \beta+\mathbf{1}, w\rangle = \langle \alpha+\beta+\mathbf{1}, w(\alpha+\beta+\mathbf{1})\rangle$, we apply **P6**. Since $\alpha \notin F$, $w(\alpha+\beta+\mathbf{1}) \subsetneq w(\alpha) \cap w(\beta+\mathbf{1})$. In particular, $\alpha \neq \beta+\mathbf{1}$, so $|w(\alpha) \cap w(\alpha+\beta+\mathbf{1})| < d$ and therefore $\alpha+\beta+\mathbf{1}$ must lie on a strictly smaller codimension face than $\beta+\mathbf{1}$. Theorem 1 then guarantees that d' must be strictly smaller than the power of log in the estimate of $I_\beta(\lambda)$ that we started with. So, the estimate in this case is strictly better because of the power of the logarithm. If equality of the inner products does not hold, the power of λ must be strictly smaller than in the estimate of $I_\beta(\lambda)$.

We move on to estimating I_{22} by applying Corollary 1.[20] We assume that $k > 1 + \langle v+\beta+\mathbf{1}, w(v+\beta+\mathbf{1})\rangle$ or else we cannot proceed because of the conditions in

[20] The major obstruction here is that $R_{m'}(x) - wx \cdot \nabla R_{m'}(x)$ does not have nice cancellation like $P_{m'}(x) - wx \cdot \nabla P_{m'}(x)$ did above (even in one variable). There is nothing we can do with this term unless ϕ is smooth convenient, or real analytic. If ϕ is smooth $m-$convenient, we can take m' to be so large that $\langle \alpha+\mathbf{1}, w(\alpha+\mathbf{1})\rangle$ is much smaller than the next term in the expansion of $I_\beta(\lambda)$ and therefore can be ignored. If ϕ is real analytic, we have to use cancellation as in the estimate of I_{21}.

Corollary 1. In the next steps, we assume β satisfies this property. By the triangle inequality and Corollary 1,

$$|I_{22}| \lesssim \max_{|v|=m'} \lambda^{-\langle v+\beta+\mathbf{1}, w(v+\beta+\mathbf{1})\rangle+1} \log^{d-1}(\lambda).$$

The vector v cannot lie on any compact face of $N(\phi)$ since we assumed $N(\phi) = N(P_m) = N(P_{m'})$ intersects each coordinate axis and $P_{m'}$ has degree at most m'. Although we don't consider the case $m' = m$, we mention that in this case, if we can conclude $\langle v, w\rangle > 1$, then the exponent of λ is strictly better. This is an assumption we need to impose if we want better estimates. Otherwise, we can only say by **P6** that if $|w(\beta + \mathbf{1})| \geq d$, then the exponent of log is smaller. If $m' > m$, for convenient phases it is easy to see that $\langle \beta + \mathbf{e}_j, w\rangle > \langle \beta, w\rangle$ for all β, j, and w, since no component of any normal w can be zero. In this case we get a better estimate. If ϕ is real analytic, then $R_{m'}(x) - wx \cdot \nabla R_{m'}(x)$ is analytic and can be expressed as $\sum_{j=1}^{d} \sum_{|\alpha| \geq m', \alpha \notin F} \alpha_i c_\alpha (v_i - w_i) x^\alpha$, just like the polynomial in I_{21}. In this case we also get a better estimate.

Let $p_0 < p_1 < \cdots$ be the well ordering stated at the beginning of the chapter. Let $p_j = \langle \beta + \mathbf{1}, w\rangle$. To summarize, we have just finished proving

$$|D_\beta I_\beta(\lambda)| \lesssim \begin{cases} \lambda^{-p_{j+1}} \log^{d_{j+1}-1}(\lambda) & d_j = 1 \\ \lambda^{-p_j} \log^{d_j-2}(\lambda) & \text{otherwise.} \end{cases}$$

7.2 Estimating higher derivatives of $I(\lambda)$

7.2.1 Analytic phases

In this section we use $\cdot$ to represent the dot product whenever a gradient is involved, e.g., $\nabla \cdot F$ is the divergence of F. In the case when ϕ is analytic, we simply write $\phi(x) = \sum_\alpha x^\alpha$ or else it is not as clear why the remainders do not contribute too much. We are free to exchange infinite sums and integrals in this part of the proof because of the Lebesgue Dominated Convergence Theorem; all the sums represent analytic functions, so are uniformly and absolutely convergent in the support of the amplitude.

Denote by $G_{n,r}(\lambda)$ the integral

$$\left(\lambda\frac{d}{d\lambda} + p_n\right)^r \left(\lambda\frac{d}{d\lambda} + p_{n-1}\right)^{d_{n-1}} \cdots \left(\lambda\frac{d}{d\lambda} + p_0\right)^{d_0} I(\lambda).$$

The induction hypothesis for analytic phases is:

$$G_{n,r}(\lambda) = \sum_{j=0}^{d_0+\cdots+d_{n-1}+r} \lambda^j J_{j,n,r}(\lambda),$$

where $J_{j,n,r}$ can be split up into finitely many integrals $J_{j,n,r,\ell}$ such that there exist

- $\beta = \beta_{n,r}$ lying in a codimension r, but not codimension $r+1$, face and a normal $w \in w(\beta+\mathbf{1})$ (depending on β) such that $p_n = \langle \beta+\mathbf{1}, w\rangle - K$ for some $K \in \mathbb{N}$,

- compactly supported $\psi_{j,n,r,\ell} \in C^\infty$,
- a countable set $\Gamma(j,n,r,\ell)$ such that $\langle \gamma + \mathbf{1}, w(\gamma + \mathbf{1}) \rangle \geq p_n + j$ for all $\gamma \in \Gamma$.

The above are such that we can write

$$J_{j,n,r,\ell} = \int e^{i\lambda\phi(x)} \sum_{\gamma \in \Gamma(j,n,r,\ell)} b_\gamma x^\gamma \psi_{j,n,r,\ell} dx$$

If the context is clear, we suppress the indices. Applying $(\lambda \frac{d}{d\lambda} + p_n)$ to each $J = J_{j,n,r,\ell}$, we write

$$\begin{aligned}\left(\lambda \frac{d}{d\lambda} + p_n\right) J &= \lambda^j \int e^{i\lambda\phi} i\lambda\phi \sum_\gamma b_\gamma x^\gamma \psi dx \\ &+ (p_n + j)\lambda^j \int e^{i\lambda\phi} \sum_{\gamma \in \Gamma} b_\gamma x^\gamma \psi dx \\ &= \lambda^j (J'(\lambda) + (p_n + j) J(\lambda))\end{aligned}$$

We first write

$$J_1' = \int e^{i\lambda\phi} i\lambda (wx \cdot \nabla\phi) \sum_\gamma b_\gamma x^\gamma \psi,$$

and

$$J_2' = \int e^{i\lambda\phi} i\lambda (\phi - wx \cdot \nabla\phi) \sum_\gamma b_\gamma x^\gamma \psi.$$

To estimate J_1', we integrate by parts:

$$J_1' = -\sum_{l=1}^{d} \int e^{i\lambda\phi} \partial^{\mathbf{e}_l} \left(\sum_\gamma b_\gamma w_l x^{\gamma + \mathbf{e}_l} \psi \right)$$

$$= \underbrace{-\int e^{i\lambda\phi}\sum_{\gamma}\langle\gamma+\mathbf{1},w\rangle b_\gamma x^\gamma\psi}_{J'_{11}} \underbrace{-\sum_{l=1}^{d}\int e^{i\lambda\phi}\sum_{\gamma} b_\gamma x^{\gamma+\mathbf{e}_l} w_l \psi'_{x_l}}_{J'_{12}}.$$

Now $\lambda^j(J'_{11}+(p_n+j)J) =$

$$\lambda^j\int e^{i\lambda\phi}\sum_{\gamma}(p_n+j-\langle\gamma+\mathbf{1},w\rangle)b_\gamma x^\gamma\psi.$$

We assumed $\langle\gamma+\mathbf{1},w(\gamma+\mathbf{1})\geq p_n+j$. If $w\notin w(\gamma+\mathbf{1})$, the estimate is strictly better because the power of λ is at least $-\langle\gamma+\mathbf{1},w\rangle+j<-p_n$.[21] Otherwise, $p_n+j-\langle\gamma+\mathbf{1},w\rangle=0$ and the sum over γ only contains exponents outside of H_w. Therefore, the power of the logarithm in the estimate is strictly better by Theorem 1. Either way, we are done estimating this integral.

In J'_{12}, we estimate each of the d summands. Theorem 1 tells us they are bounded by $\lambda^{-\langle\gamma+\mathbf{e}_l+\mathbf{1},w(\gamma+\mathbf{e}_l+\mathbf{1})\rangle}\log^{d'-1}(\lambda)$ for d' as in Theorem 1. By property P4, either the power of log must be smaller or else $\gamma+\mathbf{e}_l+\mathbf{1}$ lies in an unbounded face. In the latter case, $w_l=0$ and we do not have to estimate the l^{th} summand. This completes the estimate of J'_{12}.

We now move on to J'_2. Write $\phi(x)=\sum_{l=1}^{d}\sum_{\alpha}\alpha_i v_i b_\alpha x^\alpha$ where each v depends on the corresponding α by the condition $\langle\alpha,v\rangle=1$. Then $\phi(x)-wx\cdot\nabla\phi(x)=$

[21]Here we could apply Taylor's theorem to the analytic function inside the integral in order to obtain a polynomial plus remainder term. This way, we don't need to worry about an infinite sum of remainders in λ.

$\sum_{l=1}^{d}\sum_{\alpha}\alpha_i(v_i - w_i)b_\alpha x^\alpha$. Let us consider what happens if there are α, γ such that

$$\langle \alpha+\gamma+\mathbf{1}, w(\alpha+\gamma+\mathbf{1})\rangle = 1 + \langle \gamma+\mathbf{1}, w(\gamma+\mathbf{1})\rangle. \tag{7.2.1}$$

Letting $v_i = w_i$ for all α such that $w(\alpha) \ni w$, we see that the α under consideration lie in a strictly larger codimension face than $\beta + \mathbf{1}$, guaranteed by P6 and the fact $w(\alpha) \subsetneq w(\beta + \mathbf{1})$. In the case that equality holds in (7.2.1), we only do not get a smaller power of λ if in addition $\langle \alpha+\gamma+\mathbf{1}, w(\alpha+\gamma+\mathbf{1})\rangle = \langle \beta+\mathbf{1}, w(\beta+\mathbf{1})\rangle - K + 1 + j$. In this case we already concluded the power of log is better by Theorem 1 and P6. The estimate for analytic phases in finally finished.

7.2.2 Convenient phases

We now prove the $m-$convenient case.[22] Recall that we assume

$$k > d(2m+d)(n+1) + d.$$

The induction hypothesis for $m-$convenient phases is similar to what we had in the previous section:

$$G_{n,r}(\lambda) = \sum_{j=0}^{d_0+\cdots+d_{n-1}+r} \lambda^j J_{j,n,r}(\lambda),$$

where $J_{j,n,r}$ can be split up into finitely many integrals $J_{j,n,r,\ell}$ such that there exist

[22]The trick here is to just have the right bounds on m' and keep the previous assumptions about the structure of $G_{n,r}$; we cannot estimate $R_{m'}$ better than trivially.

- $\beta = \beta_{n,r}$ lying in a codimension r but not codimension $r+1$ face and a normal $w \in w(\beta + \mathbf{1})$ (depending on β) such that $p_n = \langle \beta + \mathbf{1}, w \rangle - K$ for some $K \in \mathbb{N}$,

- compactly supported $\psi_{j,n,r,\ell} \in C^{k-(d_0+\cdots+d_{n-1}+r)}$,

- a finite set $\{\sigma^j\}_{1 \le j \le L} \subset \mathbb{N}^d$,

- and a finite set Γ such that for all $\gamma \in \Gamma$,

$$(d_0 + \cdots + d_{n-1} + r)(m+n) + d \ge \langle \gamma + \mathbf{1}, w(\gamma + \mathbf{1}) \rangle \ge p_n + j - L,$$

  where γ satisfies $\gamma = \alpha^1 + \cdots + \alpha^{j-L}$ for some $\alpha^i \in N(\phi)$.

We assume that under the above conditions, the integrals $J_{j,n,r}$ can be written as a sum of

$$J_{j,n,r,\ell}(\lambda) = \int_{\mathbb{R}^d} e^{i\lambda\phi} (x^{\sigma^1} \partial^{\sigma^1} R_{m'}) \cdots (x^{\sigma^L} \partial^{\sigma^L} R_{m'}) \sum_{\gamma} b_\gamma x^\gamma \psi_{j,n,r,\ell}.$$

Theorem 1 showed the $n = 0$ and $r = 0$ case, taking $\psi_{0,0,0,0} = \psi$, $R_{0,0,0,0} = 0$, $\beta = \mathbf{0}$, and $\Gamma = \{\mathbf{0}\}$ with $b_{0,0,0,0} = 1$. The case $n = 0$, $r = 1$ was shown in the previous section:

$$J_{0,0,1} = -\sum_{w_m \neq 0} \int e^{i\lambda\phi(x)} x_m w_m \psi'_{x_m}(x) dx + \int e^{i\lambda\phi(x)} \sum_{\gamma \notin F} b_\gamma x^\gamma \psi(x) dx$$

$$+ \int e^{i\lambda\phi(x)} (R_m - wx \cdot \nabla R_m) \psi(x) dx.$$

From now on the dependence on j, r, n, ℓ is suppressed.

Assuming $0 \le r \le d_n - 1$, we apply $(\lambda \frac{d}{d\lambda} + p_n)$ to $G_{n,r}$. By induction hypothesis, $G_{n,r}$ is a sum of terms $\lambda^j J(\lambda)$ satisfying the conditions above, where J depends on j, r, n, and ℓ. For all j there is some β and w corresponding to a codimension 1 face such that $p_n = \langle \beta + \mathbf{1}, w \rangle - K$. Therefore,

$$\left(\lambda \frac{d}{d\lambda} + p_n\right) \lambda^j J(\lambda) = \lambda^j \Bigg(\sum_{\ell} \int e^{i\lambda\phi} i\lambda\phi (x^{\sigma^1} \partial^{\sigma^1} R_m) \cdots (x^{\sigma^L} \partial^{\sigma^L} R_m) \sum_{\gamma} b_\gamma x^\gamma \psi$$
$$+ (p_n + j) \sum_{\ell} \int e^{i\lambda\phi} (x^{\sigma^1} \partial^{\sigma^1} R_m) \cdots (x^{\sigma^L} \partial^{\sigma^L} R_m) \sum_{\gamma} b_\gamma x^\gamma \psi dx \Bigg)$$
$$= \lambda^j \Big(J'(\lambda) + (p_n + j) J(\lambda) \Big).$$

We separately estimate each summand of J' and J, and show the derivative is of the form we claimed in the induction hypothesis. For simplicity, write each summand of J' as $J'_1 + J'_2$, where

$$J'_1 = \int e^{i\lambda\phi} i\lambda (wx \cdot \nabla\phi) (x^{\sigma^1} \partial^{\sigma^1} R_m) \cdots (x^{\sigma^L} \partial^{\sigma^L} R_m) \sum_{\gamma} b_\gamma x^\gamma \psi,$$

and

$$J'_2 = \int e^{i\lambda\phi} i\lambda (\phi - wx \cdot \nabla\phi) (x^{\sigma^1} \partial^{\sigma^1} R_m) \cdots (x^{\sigma^L} \partial^{\sigma^L} R_m) \sum_{\gamma} b_\gamma x^\gamma \psi.$$

Integration by parts lets us rewrite J_1' as

$$J_1' = -\sum_{l=1}^{d} \int e^{i\lambda\phi(x)} \partial^{\mathbf{e}_l} \left((x^{\sigma^1} \partial^{\sigma^1} R_m) \cdots (x^{\sigma^L} \partial^{\sigma^L} R_m) \sum_\gamma b_\gamma w_l x^{\gamma+\mathbf{e}_l} \psi \right).$$

Now we perform some bookkeeping on where the derivatives can land inside J_1'. First, we discuss when we can apply Corollary 1. We first bound $\max |\sigma^l| + \langle \sigma^1 + \cdots + \sigma^L + \gamma + \mathbf{1}, w(\sigma^1 + \cdots + \mathbf{1})\rangle$. The terms $|\sigma^l|$ are bounded above by $d_0 + \cdots + r + 1 \leq d_0 + \cdots + d_n$ because that is the largest amount of derivatives we applied to $I(\lambda)$. Next, each $\langle \sigma^l, w(\sigma^1 + \cdots + \mathbf{1})\rangle \leq m'$ because all w satisfy $w_i \leq 1$.[23] Next, we know $\langle \gamma + \mathbf{1}, w(\gamma + \mathbf{1})\rangle \leq (d_0 + \cdots + d_n)m' + d$. Hence, the entire quantity is bounded above by $(d_0 + \cdots + d_n)(2m + 2n + 1) + d < k$.

(1) If each partial derivative $\partial^{\mathbf{e}_l}$ lands on $\sum_\gamma b_\gamma x^{\gamma+\mathbf{e}_l}$, after summing over l and adding the integral $(p_n + j)J(\lambda)$, we can abuse the following cancellation:

$$\sum_\gamma (p_n + j - \langle \gamma + \mathbf{1}, w\rangle) \int e^{i\lambda\phi} (x^{\sigma^1} \partial^{\sigma^1} R_m) \cdots (x^{\sigma^L} \partial^{\sigma^L} R_m) \sum_\gamma b_\gamma x^\gamma \psi dx.$$

If $L \neq 0$, each term inside the sum over γ can be bounded by Corollary 1. The power of λ is

$$-\langle \sigma^1 + \cdots + \sigma^L + \gamma + \mathbf{1}, w(\sigma^1 + \cdots + \mathbf{1})\rangle > -\langle \gamma + \mathbf{1}, w(\gamma + \mathbf{1})\rangle + p_n + j$$

[23] Every Newton polyhedron under consideration is contained in the Newton polyhedron of $x_1 + \cdots + x_d$. In particular, $\langle \mathbf{e}_i, w\rangle \leq 1$ for all $1 \leq i \leq d$, so that each w normal corresponding to a codimension 1 face satisfies $w_i \leq 1$.

by induction hypothesis on $\gamma \in \Gamma$ and because $\langle \sigma^1, w(\sigma^1) \rangle > 1$ for $m-$convenient phases whenever $m' > m$.[24] We now assume $L = 0$. In this case, the argument is exactly the same as for analytic functions in the previous section. However, in this section we shall be very careful about the induction and smoothness. In order to apply Corollary 1, the smoothness condition needs to be met; in the remarks above we showed indeed it is.

(2) If a partial derivative $\partial^{\mathbf{e}_l}$ lands on ψ, we see that $\langle \gamma + \mathbf{e}_l + \mathbf{1}, w(\gamma + \mathbf{e}_l + \mathbf{1}) \rangle \geq p_n - L$ and equality holds iff the power of log gives a better estimate, by the induction hypothesis on $\Gamma \ni \gamma$. If equality does not hold, the inner product is greater than $p_{n+1} - L$. Again, this form for the integral matches the statement in the induction hypothesis.[25] In this case, the inner product bounds on the new set Γ' also hold.

(3) If a partial derivative lands on one of the remainder terms,

$$\partial^{\mathbf{e}_l}(x^{\sigma+\mathbf{e}_l}\partial^{\sigma} R_m) = (\sigma + 1)x^{\sigma}\partial^{\sigma} R_m + x^{\sigma+\mathbf{e}_l}\partial^{\sigma+\mathbf{e}_l} R_m.$$

In this harder case we also apply Corollary 1. We require the same bound on k since $r \leq d_n$. Everything in the induction still holds.

[24] One main reason we had a separate section for analytic phases is because this argument would not hold if σ^i was contained in an unbounded face not contained in a coordinate hyperplane.

[25] The exponent may not have improved, but we have the same power of λ multiplying the integral as we started with.

Next, we need to consider the integral J_2'. We handle this integral as before, breaking up $\phi - wx \cdot \nabla\phi$ into $P_m - wx \cdot \nabla P_m = \tilde{P}$ and $R_m - wx \cdot \nabla R_m = \tilde{R}$.

For the integral of $\tilde{P}$, we use the induction hypothesis to bound the exponent of λ just as in the analytic case. If $L > 0$, we can below by

$$
\begin{aligned}
& j - \langle \sigma^1 + \cdots + \sigma^L + \alpha + \gamma + \mathbf{1}, w(\sigma^1 + \cdots + \mathbf{1}) \rangle \\
& \leq j - \langle \sigma^1, w(\sigma^1) \rangle - \cdots - \langle \sigma^L, w(\sigma^L) \rangle - 1 - \langle \gamma + \mathbf{1}, w(\gamma + \mathbf{1}) \rangle \\
& < j - L - 1 - p_n - j + L + 1 = -p_n.
\end{aligned}
$$

In this case we get a higher exponent of λ. If $L = 0$, then we argue as before if $p_n + j = \langle \gamma + \mathbf{1}, w(\gamma + \mathbf{1}) \rangle$, the power of log gives a better estimate.

To estimate the remainders, we note that $k > m' = m + (n+1)d$, so we can keep differentiating. In particular, any monomial x^α appearing in the remainder satisfies $\langle \alpha + \mathbf{1}, w \rangle > 1$ because $m + (n+1)d > m$.

This completes the induction proof of (7.1.1).

7.3 A differential inequality

The last thing we need to do is show (7.1.1) implies the asymptotic expansion of $I(\lambda)$ we have been trying to prove. The expansion is a corollary of the following result:

Lemma 3. *Let $f : (2, \infty) \to \mathbb{C}$ be smooth. Assume there are positive rationals*

$p_0 < p_1 < \cdots < p_{n+1}$ and positive integers $d_0, \cdots, d_{n+1}$ such that

$$\left|\left(\lambda\frac{d}{d\lambda} + p_n\right)^{d_n} \cdots \left(\lambda\frac{d}{d\lambda} + p_0\right)^{d_0} f(\lambda)\right| \lesssim \lambda^{-p_{n+1}} \log^{d_{n+1}-1}(\lambda). \tag{7.3.1}$$

Then, there are constants $a_{j,r} \in \mathbb{C}$ such that

$$f(\lambda) = \sum_{j=0}^{n} \sum_{r=0}^{d_n-1} a_{j,r} \lambda^{-p_j} \log^{d_j-1-r}(\lambda) + O(\lambda^{-p_{n+1}} \log^{d_{n+1}-1}(\lambda)).$$

First, we require some more basic results about the differential operator we are considering. We let p_j and d_j as in Lemma 3. We assume $f : \mathbb{R} \to \mathbb{C}$ is smooth for all statements below. Also, big-O statements are for $\lambda \to \infty$.

Proposition 7. *Let $h : (2, \infty) \to \mathbb{C}$ be smooth. Assume there are positive rationals $p_0 < p_1 < \cdots < p_n$ and positive integers $d_0, \cdots, d_n$ such that*

$$\left(\lambda\frac{d}{d\lambda} + p_n\right)^{d_n} \cdots \left(\lambda\frac{d}{d\lambda} + p_0\right)^{d_0} h(\lambda) = 0.$$

Then, there are $a_{j,r} \in \mathbb{C}$ such that

$$h(\lambda) = \sum_{j=0}^{n} \sum_{r=0}^{d_j-1} a_{j,r} \lambda^{-p_j} \log^{r}(\lambda).$$

Proposition 7 can be shown by an induction argument on $0 \leq m \leq d_0 + \cdots + d_n$.

Proof. First, if $p > 1$, we can integrate by parts to see

$$\int_t^\infty \lambda^{-p} \log^n(\lambda) = \sum_{j=0}^{n} c_j t^{-p+1} \log^j(t).$$

Next, we show $(\lambda\frac{d}{d\lambda} + p)(\lambda\frac{d}{d\lambda} + q)h = (\lambda\frac{d}{d\lambda} + q)(\lambda\frac{d}{d\lambda} + p)h$. Simply expanding both sides, we see they are both equal to

$$\lambda^2\frac{d^2h}{d\lambda^2} + (p+q+1)\lambda\frac{dh}{d\lambda} + pq.$$

We now show the base case of our induction. Assume $(\lambda\frac{d}{d\lambda} + p)h = 0$. Then $0 = \lambda^{p-1}(\lambda\frac{d}{d\lambda} + p)h = \frac{d}{d\lambda}(\lambda^p h)$. Integrating both sides, we see that $h(\lambda) = c_1\lambda^{-p}$.

Assume $(\lambda\frac{d}{d\lambda} + p_n)^{d'_n} \cdots (\lambda\frac{d}{d\lambda} + p_0)^{d'_0}h = 0$. The induction hypothesis states that

$$h(\lambda) = \sum_{j=0}^{n}\sum_{r=0}^{d'_j-1} a_{jk}\lambda^{-p_j}\log^r(\lambda).$$

Now we solve

$$\left(\lambda\frac{d}{d\lambda} + p\right)\left(\lambda\frac{d}{d\lambda} + p_n\right)^{d'_n} \cdots \left(\lambda\frac{d}{d\lambda} + p_0\right)^{d'_0} h = 0.$$

Since each operator commutes, write $(\lambda\frac{d}{d\lambda} + p_n)^{d'_n} \cdots (\lambda\frac{d}{d\lambda} + p_0)^{d'_0}(\lambda\frac{d}{d\lambda} + p)h = 0$. By induction,

$$\left(\lambda\frac{d}{d\lambda}+p\right)h=\sum_{j=0}^{n}\sum_{r=0}^{d'_j-1}a_{j,r}\lambda^{-p_j}\log^r(\lambda).$$

Therefore, as in the base case,

$$\frac{d}{d\lambda}(\lambda^p h)=\sum_{j=0}^{n}\sum_{r=0}^{d'_j-1}a_{j,r}\lambda^{p-p_j-1}\log^r(\lambda).$$

We integrate both sides to obtain

$$\lambda^p h(\lambda)=\begin{cases}\sum_{j\neq i}^{n}\sum_{r=0}^{d'_j-1}a'_{j,r}\lambda^{p-p_j}\log^r(\lambda)+\sum_{r=1}^{d'_i}a'_{i,r}\log^r(\lambda)+C & p=p_i\\ \sum_{j=0}^{n}\sum_{r=0}^{d'_j-1}a'_{j,r}\lambda^{p-p_j}\log^r(\lambda)+C & p\neq p_i\end{cases}$$

So if $p=p_i$, we get an additional summand $a'_{i,d'_i}\lambda^{-p_i}\log^{d'_i}(\lambda)$. Otherwise, we get an extra summand $C\lambda^{-p}$. This completes the proof. □

Proposition 8. *Let* $f:(2,\infty)\to\mathbb{C}$ *be smooth. Let* $0<p<q$ *and let* $d\in\mathbb{N}$. *If* $|(\lambda\frac{d}{d\lambda}+p)f(\lambda)|\lesssim\lambda^{-q}\log^d(\lambda)$, *then* $|f(\lambda)|\lesssim\lambda^{-q}\log^d(\lambda)$.

Proof. We multiply both sides of the inequality by λ^{p-1}, notice the left-hand side becomes exact, and integrate:

$$\left|\int_\lambda^\infty\frac{d}{dt}(t^pf(t))dt\right|\leq\int_\lambda^\infty\left|\frac{d}{dt}(t^pf(t))\right|dt\lesssim\int_\lambda^\infty t^{p-q-1}\log^d(t)dt.$$

Since $p-q-1<-1$, the rightmost side is integrable, therefore so is the leftmost.

Integrating by parts, (differentiating the log term if $d \neq 0$), we conclude

$$|\lambda^p f(\lambda)| \lesssim \lambda^{p-q} \log^d(\lambda).$$

□

Proposition 8 provides the base case for the proof of Lemma 3:

Proof. Let D_n be the differential operator $(\lambda \frac{d}{d\lambda} + p_n)^{d_n} \cdots (\lambda \frac{d}{d\lambda} + p_0)^{d_0}$. Let h be the general solution to the homogeneous equation $D_n(h) = 0$ guaranteed by Proposition 7. Then to solve for f in the differential inequality (7.3.1), we need to solve $|D_n(f+h)| \lesssim \lambda^{-p_{n+1}} \log^{d_{n+1}-1}(\lambda)$. We use induction the same way as in the proof of Proposition 8, making use of $p_0 < \cdots < p_n < p_{n+1}$. We conclude

$$|f(\lambda) + h(\lambda)| \lesssim \lambda^{-p_{n+1}} \log^{d_{n+1}-1}(\lambda).$$

Hence, there are constants $a_{j,r} \in \mathbb{C}$ such that

$$f(\lambda) = \sum_{j=0}^{n} \sum_{r=0}^{d_j-1} a_{j,r} \lambda^{-p_j} \log^r(\lambda) + O(\lambda^{-p_{n+1}} \log^{d_{n+1}-1}(\lambda)).$$

□

Now we can conclude that for all $n \in \mathbb{N}$, there are $a_{j,r} \in \mathbb{C}$ such that

$$\left| I(\lambda) - \sum_{j=0}^{n} \sum_{r=0}^{d_j-1} a_{j,r} \lambda^{-p_j} \log^{d_j-1-r}(\lambda) \right| \lesssim \lambda^{-p_{n+1}} \log^{d_{n+1}-1}(\lambda).$$

Finally, taking p_j and d_j as in Theorem 2, the proof is complete.

Chapter 8

Future work

Varchenko[18] did a great service finding counterexamples to some of Arnol'd's conjectures about stability of estimates.[26] There have been relatively little progress in this direction since Varchenko's negative results 40 years ago. Karpushkin has published some important articles studying stability of estimates, e.g., [10], [11], [12] using techniques from algebraic geometry. There has also been wonderful progress by Phong-Stein-Sturm[16] and Greenblatt[6] with more analytic techniques. My goal is to study the stability of oscillatory integrals

$$I(\lambda, t) = \int_{\mathbb{R}^d} e^{i\lambda\phi(x,t)}\psi(x)dx$$

in the real parameter t for suitable phases ϕ. Some difficulties arise in the current arguments because integration by parts is only in x. However, with the methods in

[26] See appendix for some of these examples.

this thesis it is already possible to study the stability if we assume something similar to nondegeneracy, e.g., if we assume for all t the phase $\phi(x,t)$ is nondegenerate in x. I believe that the methods presented here are well-suited for studying such problems. In particular, dyadically decomposing the domain and estimating each piece is a great strategy to attack such problems, e.g., assume we have a singularity at the origin and we perturb it so that each singularity is normal crossings. It seems likely that dyadically decomposing near the origin and applying Theorem 1 should be a good strategy for approaching this problem.

Another possibility is to assume more general conditions than nondegeneracy. By Hironaka's famous theorem on resolution of singularities[8] guarantees there are finitely many changes of variables one can perform so that the singularity becomes normal crossings. Hironaka's result is very abstract and is not practical for analysts. However, one could try blowing down a nondegenerate phase in order to get a degenerate phase, try to estimate the new phase, then generalize the method to new varieties. The difficulties in recycling our methods for more general singularities start at Lemma 2. We required larger and larger boxes to not contain any singularities of the polynomial components of $x\nabla\phi_F(x)$. I am not sure how to generalize this to arbitrary singularities, although it can be done for intersections of finitely many hyperplanes containing the origin.

Lechao Xiao has suggested generalizing the work of Phong-Stein-Sturm under a condition similar to nondegeneracy (consider instead phases satisfying for all x there are $1 \leq i \neq j \leq d$ such that $x_i x_j \phi''_{x_i x_j}(x) \neq 0$ away from coordinate hyperplanes).

This is a joint work in progress. The result is very elegant.

Appendix: counterexamples of Varchenko

Varchenko's article had a few counterexamples to stability of estimates and the necessity of assuming nondegeneracy. Even though he introduced the following examples in the reverse order, we first focus on a polynomial P with Newton distance less than one such that the estimate of $I(\lambda)$ with phase P is strictly better than $\lambda^{-1/t}$. We then show the estimate obtained is predicted by some term in the asymptotic expansion of $I(\lambda)$. Next, we discuss a polynomial Q such that $x\nabla Q_F(x)$ is not necessarily 0 away from coordinate axes for x small enough.

A polynomial with Newton distance less than 1

Define a polynomial P in five variables by

$$P(x) = x_4^2 + x_1^{4p} + x_2^{4p} + x_3^{4p} + x_5(x_4 - (x_1^2 + x_1^4 + x_2^2 + x_3^2)),$$

where p is an integer greater than 1. Varchenko checked for us that P is nondegenerate. After distributing x_5 inside the parentheses, we see that the Newton polyhedron has the following extreme points coming from each monomial:

(1) $v_1 = (4p, 0, 0, 0, 0)$ from x_1^{4p};

(2) $v_2 = (0, 4p, 0, 0, 0)$ from x_2^{4p};

(3) $v_3 = (0, 0, 4p, 0, 0)$ from x_3^{4p};

(4) $v_4 = (2, 0, 0, 0, 1)$ from $x_1^2 x_5$;

(5) $v_5 = (0, 2, 0, 0, 1)$ from $x_2^2 x_5$;

(6) $v_6 = (0, 0, 2, 0, 1)$ from $x_3^2 x_5$;

(7) $v_7 = (0, 0, 0, 2, 0)$ from x_4^2;

(8) $v_8 = (0, 0, 0, 1, 1)$ from $x_4 x_5$.

Note that $x_1^4 x_5$ lies inside the orthant generated by $x_1^2 x_5$, so it cannot be an extreme point. These extreme points give rise to two compact codimension 1 faces:

(1) F_1 generated by v_1, v_4, v_5, v_6, and v_7 with normal

$$w^1 = \left(\frac{1}{4p}, \frac{1}{4p}, \frac{1}{4p}, \frac{1}{2}, 1 - \frac{1}{2p}\right).$$

Note that this face also contains v_2, v_3, and notice that v_8 lies above the hyperplane with normal w^1.

(2) F_2 generated by v_4 through v_8 with normal

$$w^2 = \Big(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{2}, \frac{1}{2}\Big).$$

Each of the vectors v_1, v_2, v_3 lie above the hyperplane with normal w^2.

Since $\langle v_i, w^j \rangle \geq 1$ must be satisfied for all i, j, we can check that these are the only two compact codimension 1 faces (e.g., if one replaces v_1 with v_4 when defining a hyperplane, one sees the hyperplane intersects the interior of $N(P)$).

Since $|w^1| = \frac{1}{4p} + \frac{3}{2} > 1$, we conclude that the Newton distance of P must be less than 1, since we know $\min |w^i| = \frac{1}{t}$ where t is the Newton distance. However, Varchenko computes that

$$\left| \int_{\mathbb{R}^5} e^{i\lambda P(x)} \psi(x) dx \right| \lesssim \lambda^{-7/4} \log^4(\lambda)$$

whenever ψ is smooth and supported close enough to the origin. Clearly $-\frac{7}{4} < -\frac{3}{2} - \frac{1}{4p}$ for all $p > 1$. Theorem 2 guarantees that $-\frac{7}{4} = \langle \beta + \mathbf{1}, w(\beta + \mathbf{1}) \rangle - n > 1/t$ for some $\beta \in \mathbb{N}^d$ and $n \in \mathbb{N}$ Let $\beta = (4p-1, 0, 0, 0, 0)$ and $n = 0$. Then $\langle \beta, w^1 \rangle = \frac{4p-1}{4p}$, so $\langle \beta + \mathbf{1}, w^1 \rangle = \frac{7}{4}$. We need to check $w^1 \in w(\beta + \mathbf{1})$. Let $\lambda_1 = \frac{2p-1}{2p}$, $\lambda_2 = \lambda_3 = \frac{1}{4p}$, $\lambda_4 = 1$, and $\lambda_5 = \frac{1}{2}$. Then

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 + \lambda_4 v_5 + \lambda_5 v_7 = \beta + \mathbf{1}.$$

Let $\eta_i = 2\lambda_i/5$. Then $\sum \eta_i = 2\sum \lambda_i/5 = 1$. We conclude $2(\beta + \mathbf{1})/5 \in F_1$, since it's a convex combination of vectors in F_1, so $w^1 \in w(\beta + \mathbf{1})$.

A polynomial not satisfying Varchenko's condition

Let Q_s be the polynomial in 3 variables with positive parameter s defined by

$$Q_s(x, y, z) = (-sx^2 + x^4 + y^2 + z^2)^2 + x^{4p} + y^{4p} + z^{4p}$$

for $p > 1$. There is a normal vector $w_1 = (1/4, 1/4, 1/4)$ corresponding to the codimension one face containing $(4, 0, 0), (0, 4, 0)$, and $(0, 0, 4)$. Since this face contains $(4/3, 4/3, 4/3)$, the Newton distance is $4/3$ and therefore $1/t = 3/4$. Let F be the codimension 2 face containing $(4, 0, 0)$, $(2, 2, 0)$ and $(0, 4, 0)$, i.e., let $Q_F(x, y, z) = s^2x^4 - 2sx^2y^2 + y^4 = (sx^2 - y^2)^2$. We can check $\nabla Q_F = 0$ if $y = \sqrt{s}x$, so Varchenko's condition is not satisfied. Varchenko proved that in this case

$$\left| \int e^{i\lambda Q}\psi \right| \gtrsim \lambda^{-1/2-\gamma(p)}$$

for some amplitude ψ and some function $\gamma(p) \to 0$ as $p \to \infty$, giving a much worse estimate than $\lambda^{-1/t} = \lambda^{-3/4}$.

Bibliography

[1] Ould M. Abderrahmane. On the Lojasiewicz exponent and Newton polyhedron. *Kodai Math. J.*, 28(1):106–110, 2005.

[2] Toshizumi Fukui and Etsuo Yoshinaga. The modified analytic trivialization of family of real analytic functions. *Invent. Math.*, 82(3):467–477, 1985.

[3] Loukas Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, third edition, 2014.

[4] Michael Greenblatt. Oscillatory integral decay, sublevel set growth, and the Newton polyhedron. *Math. Ann.*, 346(4):857–895, 2010.

[5] Michael Greenblatt. Resolution of singularities, asymptotic expansions of integrals and related phenomena. *J. Anal. Math.*, 111:221–245, 2010.

[6] Michael Greenblatt. Resolution of singularities in two dimensions and the stability of integrals. *Adv. Math.*, 226(2):1772–1802, 2011.

[7] Michael Greenblatt. Maximal averages over hypersurfaces and the Newton polyhedron. *J. Funct. Anal.*, 262(5):2314–2348, 2012.

[8] Heisuke Hironaka. Resolution of singularities of an algebraic variety over a field of characteristic zero. I, II. *Ann. of Math. (2)* **79** *(1964), 109–203; ibid. (2)*, 79:205–326, 1964.

[9] Joe Kamimoto and Toshihiro Nose. Newton polyhedra and weighted oscillatory integrals with smooth phases. *Trans. Amer. Math. Soc.*, 368(8):5301–5361, 2016.

[10] V. N. Karpushkin. Uniform estimates for oscillatory integrals with a phase of the series $\tilde{R}_m$. *Mat. Zametki*, 64(3):468–469, 1998.

[11] V. N. Karpushkin. Uniform estimates for an oscillatory integral and volume with the phase of A. N. Varchenko. *Mat. Zametki*, 72(5):688–692, 2002.

[12] V. N. Karpushkin. Four theorems on uniform estimates of oscillatory integrals. *Moscow Univ. Math. Bull.*, 69(3):128–131, 2014. Translation of Vestnik Moskov. Univ. Ser. I Mat. Mekh. **2**014, no. 3, 56–60.

[13] Steven G. Krantz and Harold R. Parks. *A primer of real analytic functions*, volume 4 of *Basler Lehrbücher [Basel Textbooks]*. Birkhäuser Verlag, Basel, 1992.

[14] S. Łojasiewicz. Sur le problème de la division. *Studia Math.*, 18:87–136, 1959.

[15] Bernard Malgrange. Intégrales asymptotiques et monodromie. *Annales Scientifiques de l'École Normale Supérieure*, 7(3):405–430, 1974.

[16] D. H. Phong, E. M. Stein, and J. A. Sturm. On the growth and stability of real-analytic functions. *Amer. J. Math.*, 121(3):519–554, 1999.

[17] C.R. M. Talbot. *Enumeration of Lines of Third Order*. Ulan Press, 1860.

[18] A. N. Varchenko. Newton polyhedra and estimation of oscillating integrals. *Funktsional'nyi Analiz i Ego Prilozheniya*, 10(3):13–38, 1976.

[19] B. A. Vasil′ev. The asymptotic behavior of exponential integrals, the Newton diagram and the classification of minima. *Funkcional. Anal. i Priložen.*, 11(3):1–11, 96, 1977.

[20] Etsuo Yoshinaga. Topologically principal part of analytic functions. *Trans. Amer. Math. Soc.*, 314(2):803–814, 1989.

[21] Günter M. Ziegler. *Lectures on polytopes*, volume 152 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.