Acta Math., 179 (1997), 105–152

The Newton polyhedron and oscillatory integral operators

by

D. H. PHONG and E. M. STEIN

Columbia University
New York, NY, U.S.A.

Princeton University
Princeton, NJ, U.S.A.

1. Introduction

The lack of suitable methods of stationary phase for both degenerate oscillatory integrals and degenerate oscillatory integral operators has been a major source of difficulties in many areas of analysis and geometry. Despite their name, degenerate phases are often generic, for example in presence of high codimension or of additional parameters, a phenomenon familiar in singularity theory. Strongest results to date include the now classic work of Varchenko [18] on decay rates for oscillatory integrals with generic analytic phases, and the relatively more recent progresses in the study of Lagrangians with Whitney folds [8], [4], [3], [9], generalized Radon transforms in the plane [10], [15], and sharp forms of the van der Corput Lemma in one dimension [11], [2].

The purpose of this paper is to establish sharp and completely general bounds for oscillatory integral operators on $L^2(\mathbf{R})$ of the form

$$(Tf)(x)=\int_{-\infty}^{\infty} e^{i\lambda S(x,y)}\chi(x,y)\,f(y)\,dy, \tag{1.1}$$

where $\chi\in C_0^\infty(\mathbf{R}^2)$ is a smooth cut-off function supported in a small neighborhood of the origin, and the phase $S(x,y)$ is real-analytic. (Besides its intrinsic interest, the decay rate of $\|T\|$ in $|\lambda|$ is closely related to the regularity of Radon transforms (see e.g. [3], [10], [15]), but we shall not elaborate on this point here.) Our main result is that the sharp bounds for $\|T\|$ as an operator on $L^2(\mathbf{R})$ are determined by the (reduced) Newton polyhedron of the phase $S(x,y)$. Remarkably, the Newton polyhedron is the notion which had been shown by Varchenko, confirming earlier hypotheses of Arnold, to control the

Research supported in part by the National Science Foundation under Grants DMS-95-05399 and DMS-94-01579.A01.

apparently unrelated decay rate for the *two-dimensional, scalar* oscillatory integrals with phase $S(x,y)$.

More precisely, let

$$S(x,y)=\sum_{p,q=0}^{\infty} c_{pq}x^p y^q$$

be the Taylor series expansion of $S(x,y)$. The *Newton polyhedron* of $S(x,y)$ at the origin is defined to be the convex hull of the union of all the northeast quadrants in $\mathbf{R}^2_{\geqslant 0}$ with corners at the points (p,q) satisfying $c_{pq}\neq 0$. The *reduced Newton polyhedron* is defined in the same way, with this time the vertices (p,q) constrained by the additional requirement that $pq\neq 0$. Equivalently, the reduced Newton polyhedron is the translate by the vector $(1,1)$ of the Newton polyhedron of $\partial^2 S/\partial x\,\partial y$ at the origin. The boundaries of the Newton polyhedra are called *Newton diagrams.* We can now define the *Newton decay rate* as

$$\delta=\min_l \delta_l \tag{1.2}$$

where the index l runs through the boundary lines of the reduced Newton diagram, and $(\delta_l^{-1},\delta_l^{-1})$ is the intersection of the line l with the line $p=q$ bisecting the first quadrant. The boundary lines which realize the minimum value δ are called the *main boundary lines.* In two dimensions, there are at most two main boundary lines, the case of two occurring exactly when the two lines and the bisecting line intersect simultaneously at a vertex of the Newton diagram.

We have then the following theorem:

THEOREM 1. *Let $S(x,y)$ be a real-analytic phase function. If the support of χ is sufficiently small, then the operator T is bounded on $L^2(\mathbf{R})$ with the bound*

$$\|T\|\leqslant C|\lambda|^{-\frac{1}{2}\delta}, \tag{1.3}$$

where δ is the Newton decay rate. The result (1.3) *is exact in the sense that if χ is not zero at the origin, then $\|T\|\geqslant c'|\lambda|^{-\frac{1}{2}\delta}$, as $|\lambda|\to\infty$, for some $c'>0$.*

It is intriguing to compare this statement with Varchenko's theorem for oscillatory integrals [1], [18]. In two dimensions, this theorem asserts that for generic analytic phase functions $S(x,y)$ (non-$\mathbf{R}$-degenerate, in Varchenko's terminology), we have the decay rate

$$\left|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i\lambda S(x,y)}\chi(x,y)\,dx\,dy\right|\leqslant C|\lambda|^{-\hat{\delta}}(\log|\lambda|)^{\varkappa},$$

where $\hat{\delta}$ is defined in the same way as δ, but with the reduced Newton diagram of $S(x,y)$ replaced by its full Newton diagram, and $\varkappa=0$ or 1, depending on whether the bisectrix

$p=q$ encounters the Newton diagram at a vertex or not. The factor $\frac{1}{2}$ in the decay rate for the operator vs. the decay rate for the oscillatory integral is easily understood on the basis of dimensionality. So is the irrelevance of pure x^m- and y^n-terms in the Taylor expansion of $S(x,y)$ in the case of the operator, since these terms are readily absorbed in the $L^2(\mathbf{R})$-norms of Tf and f respectively. However, we do not have a ready explanation for the absence of logarithmic terms, nor for the full generality of the analytic phase in the operator case.

The above theorem is a substantial generalization and strengthening of both the homogeneous polynomial case of [13], and the case of analytic $S(x,y)$ satisfying

$$\frac{\partial^m S}{\partial x^{m-1}\partial y}(0)\neq 0, \quad \frac{\partial^n S}{\partial x\,\partial y^{n-1}}(0)\neq 0 \tag{1.4}$$

of [12]. Indeed, the various decay rates for $\|T\|$ found in [13] for polynomial phase functions of the form $S(x,y)=\sum_{p=1}^{n} a_p x^p y^{n-p}$ just correspond to the different values that δ can take, depending on which boundary line of the reduced Newton diagram for $S(x,y)$ is the main boundary line. For general analytic phase functions $S(x,y)$ satisfying (1.4), the estimate obtained in [12] was

$$\|T\|=O(|\lambda|^{-\frac{1}{2}\cdot\frac{m+n-4}{mn-m-n}}). \tag{1.5}$$

We note that in this case $(m-1,1)$ and $(1,n-1)$ are two extreme points of the reduced Newton diagram, and that $((m+n-4)/(mn-m-n))^{-1}$ is just the p-coordinate (or q-coordinate) of the intersection of the bisectrix $p=q$ with the line joining these two extreme points. By the convexity of the reduced Newton diagram, we have then $(m+n-4)/(mn-m-n)\leqslant\delta$, with strict inequality except in the non-generic case where the line joining $(m-1,1)$ and $(1,n-1)$ is actually a boundary line for the reduced Newton diagram. The case of smooth phases, as well as the closely related Radon transforms in the plane, is treated extensively in Seeger [15]. We also observe that when either m or n is equal to 2 in (1.4), the decay rate becomes $O(|\lambda|^{-\frac{1}{2}\cdot\frac{m+n-4}{mn-m-n}})=O(|\lambda|^{-\frac{1}{2}})$. This is of course just a special case of the classic result $\|T\|=O(|\lambda|^{-\frac{1}{2}d})$, valid for oscillatory integral operators on $L^2(\mathbf{R}^d)$ with phase function satisfying the non-degeneracy condition $\det S''_{x_j y_k}\neq 0$ [5].

Although Theorem 1 establishes a strong correlation between the decay rates for scalar and for operator oscillatory integrals, the methods of proof are completely different. Varchenko's proof in the scalar case is based on successive blow-ups of the phase which reduce the integral to some simple canonical models, for which the desired estimate can be established directly. It is not known whether such methods can be developed for the operator case. Rather, a fundamental tool we rely on in this paper is a decomposition of the complement $\mathbf{R}^2\backslash Z$ in $\mathbf{R}^2$ of the singular variety

$$Z=\{\partial_x\partial_y S\equiv S''(x,y)=0\}$$

into small curved rectangles, whose contributions are summed by balancing an "oscillatory" estimate governed by the size of $S''(x,y)$ on the box, with a "size" estimate governed by the dimensions of the box. In its simplest form, this method had been instrumental in the earlier work [10], [13]. Here we must use it in conjunction with a delicate resolution of $\mathbf{R}^2\backslash Z$, where the regions between two highly tangent branches of Z have to be suitably magnified in several successive scales.

The paper is organized as follows. In §2, we provide a few technical lemmas, which show that for our purposes, analytic functions are as well behaved as polynomials. In §3, we establish an *Operator van der Corput* Lemma, which gives decay rates for oscillatory integral operators when the Hessian is of constant size and the shape of the support does not oscillate wildly. Due to the fact that the variety Z can be extremely complicated, we give in §4 a detailed treatment of some typical cases, which serve as models for more general situations. This section is crucial to the understanding of the paper. It is actually a prerequisite for §5, since some of the key ideas and resummation techniques are introduced and discussed at length there. In §5, the proof for the general case is given, after a suitable discussion of the parameters of the Newton diagram, and of the algorithm for how to classify roots and study them with increasingly higher resolutions. Finally, we note that, in preparation for higher-dimensional cases, it may be very useful to bring to bear directly the techniques of algebraic geometry, which we have in essence circumvented here by use of Puiseux series. Some progress in this direction is in [7].

2. Functions of polynomial type

Let $F\in C^{(N)}[\alpha,\beta]$, with $N\geqslant 1$. We say that F is of polynomial type (of degree N) in $[\alpha,\beta]$ with constant C if

$$\sup_{\alpha\leqslant x\leqslant\beta}|F^{(N)}(x)|\leqslant C\inf_{\alpha\leqslant x\leqslant\beta}|F^{(N)}(x)|. \tag{2.1}$$

We always assume that N is greater than or equal to 1.

LEMMA 1. *Under the above assumption, we have*

(i) $\sup_{x\in I^*}|F(x)|\leqslant a\sup_{x\in I}|F(x)|$,

(ii) $\sup_{x\in I^*}|F'(x)|\leqslant a\delta^{-1}\sup_{x\in I}|F(x)|$,

where I is any interval in $[\alpha,\beta]$ of length δ, I^* is its double (in $[\alpha,\beta]$), and the constant a depends only on N and the constant C appearing in (2.1).

LEMMA 2. *Assume that F is in $C^{(N)}[\alpha,\beta]$, and that $|F^{(N)}(x)|\geqslant A$ in $[\alpha,\beta]$. Then for all $\mu>0$,*

$$|\{x:|F(x)|\leqslant\mu\}|\leqslant C_N\mu^{1/N}A^{-1/N}. \tag{2.2}$$

Proof of Lemma 2. The case $N=1$ is clear. We proceed by induction on N. Since $F^{(N-1)}$ is monotone, we may assume that its minimum is at an end point, say $x_0=\alpha$. The measure of the interval $[x_0,x_0+\Delta]\cap[\alpha,\beta]$ is at most Δ, while in its complement, we have $|F^{(N-1)}(x)|\geqslant A\Delta$. Now the inductive hypothesis to F applied to each of the components of the complement gives

$$|\{x\in I\backslash[x_0,x_0+\Delta]:|F(x)|\leqslant\mu\}|\leqslant 2C_{N-1}\mu^{1/(N-1)}(A\Delta)^{-1/(N-1)}.$$

Hence

$$|\{x:|F(x)|\leqslant\mu\}|\leqslant\Delta+2C_{N-1}\mu^{1/(N-1)}(A\Delta)^{-1/(N-1)}.$$

If we choose Δ to be $\Delta=\mu^{1/N}A^{-1/N}$, the inequality (2.2) follows with $C_N=1+2C_{N-1}$.

Proof of Lemma 1. By normalization, we can always assume that

$$1\leqslant|F^{(N)}(x)|\leqslant C \tag{2.3}$$

in $[\alpha,\beta]$. Consider first the case $N=1$. Then by (2.3) we have $\sup_{x\in I^*}|F'(x)|\leqslant C$. If $\mu=\sup_{x\in I}|F(x)|$, the lemma implies that $|I|=\delta\leqslant c_1\mu$. Hence $\sup_{x\in I^*}|F'(x)|\leqslant C\leqslant a\delta^{-1}\mu=a\delta^{-1}\sup_{x\in I}|F(x)|$ if $a=c_1C$, and (ii) is proved in this case. Next, if $x_1\in I^*$ and $x_2\in I$, then

$$|F(x_1)-F(x_2)|\leqslant|x_1-x_2|\sup_{x\in I^*}|F'(x)|\leqslant 2\delta\cdot a\delta^{-1}\sup_{x\in I}|F(x)|.$$

Thus $\sup_{x\in I^*}|F(x)|\leqslant\sup_{x\in I}|F(x)|+2a\sup_{x\in I}|F(x)|$, and (i) is proved for $N=1$ (with a replaced by $1+2a$).

We can now prove Lemma 1 for general N by induction on N. Assume then that it holds for $N-1$, and set $\sup_{x\in I}|F(x)|=\mu$. In view of Lemma 2, $|I|=\delta\leqslant C_N\mu^{1/N}$, and thus $\delta^N\leqslant C_N^N\mu$. For x, $x+\bar h\in I$, we have

$$F(x+\bar h)=\sum_{k=1}^{N-1}\frac{\bar h^k}{k!}F^{(k)}(x)+O(\bar h^N)\sup_{x\in I}|F^{(N)}(x)|. \tag{2.4}$$

Next, for any $x\in I$, there is an h, $|h|=\delta\Delta^{-1}=\frac12|I|$, so that $x+sh\in I$ for all $0\leqslant s\leqslant 1$. If we integrate (2.4) with $\bar h=sh$, we find

$$\int_0^1 F(x+sh)\psi(s)\,ds=\sum_{k=1}^{N-1}\frac{h^k}{k!}F^{(k)}(x)\int_0^1 s^k\psi(s)\,ds+O(\delta^N).$$

Choose ψ so that $\int_0^1 s\psi(s)\,ds=1$ and $\int_0^1 s^k\psi(s)\,ds=0$ for $k=0,2,3,\ldots,N-1$. Since

$$\left|\int_0^1 F(x+sh)\,\psi(s)\,ds\right| \leqslant C\sup_{x\in I}|F(x)| = c\mu$$

we obtain $\frac{1}{2}\delta|F'(x)|\leqslant c(\mu+\delta^N)\leqslant C\mu$. Hence

$$\sup_{x\in I}|F'(x)| \leqslant c'\mu\delta^{-1}. \tag{2.5}$$

We can now apply the case $N-1$ to the function F'. In view of part (i), the inequality (2.5) implies that

$$\sup_{x\in I^*}|F'(x)| \leqslant ac'\mu\delta^{-1} = ac'\delta^{-1}\sup_{x\in I}|F(x)|. \tag{2.6}$$

This is the statement (ii) for the case N. From (2.6), we deduce that $\sup_{x\in I^*}|F(x)|\leqslant a\sup_{x\in I}|F(x)|$, as in the case $N=1$, and the proof of Lemma 1 is complete.

3. Puiseux expansions; beginning of the proof

The proof of the theorem is based on a decomposition of the operator T into pieces supported on certain curved boxes, away from the singular variety where the Hessian of the phase $S(x,y)$ vanishes. A key ingredient is the following operator version of the van der Corput Lemma, adapted to curved boxes (a version with rectangles was instrumental in the study of homogeneous polynomial phases in [13]).

THE OPERATOR VAN DER CORPUT LEMMA. *Let ϕ be a monotone continuous function on $[\alpha,\beta]$. Let*

$$\mathcal{B} = \{(x,y) : \phi(x) < y < \phi(x)+\delta,\ \alpha\leqslant x\leqslant\beta\} \tag{3.1}$$

denote the "curved box" of thickness δ defined by ϕ. Assume that $\chi(x,y)$ is supported in $\mathcal{B}$, and that

(i) $|\partial_y^n\chi(x,y)|\leqslant C\delta^{-n}$, $n=0,1,2$;

(ii) *the function $S''(x,y)$ is of polynomial type of order N in y, uniformly in x, in the "double" $\mathcal{B}^*$ defined by*

$$\mathcal{B}^* = \{(x,y) : \phi(x)-\delta < y < \phi(x)+2\delta,\ \alpha\leqslant x\leqslant\beta\};$$

(iii) $\mu\leqslant|S''(x,y)|\leqslant A\mu$, *for some positive constants A and μ.*

Then

$$\|T\| \leqslant C_{A,N}(\lambda\mu)^{-1/2} \tag{3.2}$$

for some constant $C_{A,N}$ independent of χ, δ and the interval $[\alpha,\beta]$.

Proof of the Operator van der Corput Lemma. We consider first the special case where ϕ is linear, i.e., $\mathcal{B}$ is a parallelogram with a pair of sides parallel to the y-axis.

This is then Lemma 1.1 in [13], except for the assumption there that $S(x,y)$ was a polynomial in y of bounded degree, as x varies. However, since S'' is assumed here to be of polynomial type, we can invoke the inequalities (i) and (ii) of Lemma 1 in §2 as a substitute for Lemma 1.2 in [13]. The proof is otherwise unchanged.

To prove the Operator van der Corput Lemma in the general case, assume that ϕ is increasing. Define $\alpha = x_0, x_1, ..., x_M = \beta$ so that

$$\phi(x_j) = j\delta + \phi(x_0), \quad 0 \leqslant j \leqslant M-1,$$
$$\phi(x_M) - \phi(x_{M-1}) \leqslant \delta.$$

Let $\mathcal{R}_j$ denote the rectangle $[x_j, x_{j+1}] \times [\phi(x_j), \phi(x_j)+2\delta]$. Then

$$\mathcal{B} \subset \bigcup_{j=0}^{M-1} \mathcal{R}_j$$

and

$$T = \sum_j T_j$$

if we set

$$T_j f(x) = \int_{\mathcal{R}_j} e^{i\lambda S(x,y)} \chi(x,y) f(y)\, dy.$$

By the special case we just proved, we have $\|T_j\| \leqslant A(\lambda\mu)^{-1/2}$. Moreover, $T_j T_k^* = T_j^* T_k = 0$ when $|j-k| \geqslant 2$, since both the x- and y-supports of the kernels of T_j and T_k are then disjoint. The Operator van der Corput Lemma follows.

Proof of Theorem 1. We decompose the operator T as

$$T = \sum_{\pm} \sum_{j,k} T_{jk}^{\pm,\pm} \tag{3.3}$$

with $T_{jk}^{\pm,\pm}$ defined by

$$T_{jk}^{\pm,\pm} f(x) = \int_{-\infty}^{\infty} e^{i\lambda S(x,y)} \chi_j(x) \chi_k(y) \chi(x,y) f(y)\, dy. \tag{3.4}$$

Here $\sum_j \chi_j(x) = 1$ is a dyadic partition of unity of $\mathbf{R}_+$, with χ_j supported in the interval $[2^{-1-j}, 2^{1-j}]$. The indices $\pm$ refer to the quadrants defined by specific signs for x and for y. To simplify our notation, we shall restrict our discussion to the northeast quadrant $x>0$, $y>0$, the others being exactly similar, and drop the upper indices $\pm$.

Let K be a large positive constant which we shall specify later. We shall sum back separately the contributions of the ranges

$$k > j-K \quad \text{and} \quad k \leqslant j-K. \tag{3.5}$$

Consider first the range $k>j-K$. In this range, we shall parametrize the singular variety where $S''=0$ in terms of x. More precisely, the analyticity of $S(x,y)$ and the Weierstrass Preparation Theorem imply that, up to a non-vanishing prefactor $U(x,y)$, we can write $S''(x,y)$ as a polynomial in y with analytic coefficients in x,

$$S''(x,y)=U(x,y)x^{\alpha}(y^n+c_{n-1}(x)y^{n-1}+\ldots+c_0(x))y^{\beta}. \tag{3.6}$$

We note that the reduced Newton diagram of $S(x,y)$ is the same as the translate by the vector $(1,1)$ of the Newton diagram of $S''(x,y)$. Furthermore, Newton diagrams are invariant under multiplication by a smooth non-vanishing function. Indeed, consider two functions $\Phi(x,y)$ and $\widetilde{\Phi}(x,y)=U(x,y)\Phi(x,y)$ related by a smooth non-vanishing factor $U(x,y)$, and their Taylor expansions

$$\Phi=\sum_{p,q=0}^{\infty}c_{pq}x^py^q, \quad \widetilde{\Phi}(x,y)=\sum_{\tilde{p},\tilde{q}=0}^{\infty}\tilde{c}_{\tilde{p}\tilde{q}}x^{\tilde{p}}y^{\tilde{q}}.$$

Then $\tilde{c}_{\tilde{p}\tilde{q}}=U(0)c_{\tilde{p}\tilde{q}}$ unless there is a coefficient $c_{pq}=0$ with $(p,q)\neq(\tilde{p},\tilde{q})$, $0\leqslant p\leqslant\tilde{p}$ and $0\leqslant q\leqslant\tilde{q}$. In particular, if $c_{\tilde{p}\tilde{q}}$ is an extreme point of the Newton diagram of Φ, there is no non-vanishing c_{pq} satisfying the preceding condition, and $\tilde{c}_{\tilde{p}\tilde{q}}$ must be non-zero. This shows that the Newton diagram of Φ must be contained in that of $\widetilde{\Phi}$, and, by reversing the roles of Φ and $\widetilde{\Phi}$, that they are actually identical.

Returning to the case at hand, we deduce that the Newton diagram of $S''(x,y)$ (or the reduced Newton diagram of $S(x,y)$) is the same as that of the polynomial

$$x^{\alpha}(y^n+c_{n-1}(x)y^{n-1}+\ldots+c_0(x))y^{\beta} \tag{3.7}$$

(or its translate by the vector $(1,1)$).

It is well known that the non-trivial zeroes $r_s(x)$, $s=1,\ldots,n$, of the polynomial of order n in y in (3.7) can be expressed in a small neighborhood of 0 as Puiseux series

$$r_s(x)=c_sx^{a_s}+\ldots \tag{3.8}$$

with the exponents a_s positive real numbers. In fact, the polynomial

$$y^n+c_{n-1}(x)y^{n-1}+\ldots+c_0(x)$$

can be first factored into a finite product of polynomials in y with analytic coefficients, with the property that no factor in the product admits by itself identical roots (see e.g. Saks and Zygmund [14, pp. 268–271]). If we consider next each of the factors, the arguments in Siegel [16, pp. 90–98] can be easily adapted to produce the desired Puiseux series expansion. More specifically, we note that, by a simple application of the Implicit Function Theorem, the roots of each of these factors are analytic functions of x in a small neighborhood of any value x_0 for which they are pairwise distinct. Assume then that at $x_0=0$, we have some multiple roots. Since after factorization, each factor has no identically equal roots, the values of x_0 where the roots can be multiple must be isolated points in the complex x-plane. In particular, there is a small pointed disk around $x=0$, where the roots are all distinct and can be given an ordering. If we analytically continue these roots around the origin, they will come back with possibly some permutation. Since the order of any permutation is finite, we can reiterate the analytic continuation a finite number N of times until we get back the original ordering. Now the x-range defined by going around the origin N times and gluing back with the original values is holomorphically equivalent to a disk in the t-plane, with $t=x^{1/N}$. By construction, the roots are analytic functions in the pointed t-disk, continuous at $t=0$, and hence must be analytic functions in the full t-disk. Since a power series in t is obviously a Puiseux series in x, this establishes our claim.

The leading exponents a_s in the roots (3.8) as well as the coefficients c_s can of course be the same although the roots are distinct. We shall in fact, in §5, introduce a more systematic notation to deal with this phenomenon. For the moment, we simply list all the *distinct* leading exponents a_l,

$$a_l < a_{l+1}, \tag{3.9}$$

and choose K large enough so that there can be no major cancellation in $y-r_s(x)$ when $a_s<1$, and the support of χ small enough so that we have when $\chi_j(x)\chi(x,y)\neq 0$,

$$a_l j+10K < a_{l+1} j. \tag{3.10}$$

Before proceeding further, we present in the next section some model cases upon which the general proof will be built.

4. Model cases

Due to the increasingly fine resolution with which we have to probe the zero set of the Hessian of the phase function, the complete argument for Theorem 1 requires a cumbersome notation and induction process. However, the main ideas can be easily illustrated in a few model cases, which we present in increasing order of generality. By the time we are done with Model V, we will essentially have dealt with the most general case.

(a) *Model* I. In this case, the operator T is given by (1.1), with a phase function $S(x,y)$ satisfying

$$S''(x,y)=(y-x^a)^n. \tag{4.1}$$

The analyticity of S'' implies that both a and n are non-negative integers. If either a or n is 0, the phase is actually non-degenerate in a small neighborhood of the origin. Thus we may always assume that $a\geqslant 1$ and $n\geqslant 1$.

Besides boundary lines parallel to the axes, the reduced Newton diagram of $S(x,y)$ admits a single boundary segment, of equation

$$q=\frac{-p+1+a+na}{a}. \tag{4.2}$$

This segment is the main segment, and the Newton decay rate δ,

$$\|T\|\leqslant C|\lambda|^{-\frac{1}{2}\delta}, \tag{4.3}$$

is given by

$$\delta=\frac{1+a}{1+a+an}. \tag{4.4}$$

On the support of the cut-off functions $\chi_k(y)\chi_j(x)$, where we restrict ourselves to considering only the quadrant $\{x>0,y>0\}$, the other quadrants being similar, we have

$$2^{-k-1}<y<2^{-k+1},$$
$$2^{-ja-a}<x^a<2^{-ja+a}.$$

Thus we need to consider three ranges of j,k, corresponding to whether there can be major cancellations between y and x^a:

- $2^{-ja+a}<2^{-k-2}$, i.e., $k<ja-a+2$;
- $2^{-k+1}<2^{-ja-a-1}$, i.e., $k>ja+a+2$;
- $ja-a+2\leqslant k\leqslant ja+a+2$.

It is convenient to denote simply the first range by $k\ll ja$, where the symbol $\ll$ means that the left-hand side is smaller than the right-hand side, even after the addition of a positive constant depending only on the exponent a and possibly on the coefficients of the Puiseux series in the factorization of S'', but not on the summation indices j and k. Similarly for $\gg$. Finally, we denote the third type of range by

$$k\sim aj.$$

We note that the number of k's in this range is bounded by a fixed constant.

We also denote by Δx the maximum length of the support of the kernel of T_{jk}, viewed as a function of x, with y as a parameter. Similarly, the maximum y-length of the support of T_{jk} will be denoted by Δy. Since the kernel of T_{jk} is uniformly bounded, we have the following basic "size estimate":

$$\|T_{jk}\| \leqslant (\Delta x)^{\frac{1}{2}} (\Delta y)^{\frac{1}{2}} \leqslant 2^{-\frac{1}{2}(j+k)}, \tag{4.5}$$

where here (as elsewhere in the paper), we have dropped all constants independent of λ, j and k.

In the range $k \ll aj$, the phase satisfies

$$S'' \sim 2^{-kn}.$$

Furthermore, the support of $\chi_j(x)\chi_k(y)$ is clearly of the admissible curved-box form, and the phase function $S(x,y)$ is of polynomial type in y. Thus the Operator van der Corput Lemma applies, and we have

$$\|T_{jk}\| \leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}kn}. \tag{4.6}$$

Our general strategy is to sum the contributions of $\|T_{jk}\|$'s by balancing the size estimate (4.5) with the "oscillatory" estimate (4.6). To enforce the constraint $k \ll aj$, we set

$$aj = k + r,$$

where r is an integer bounded from below. Then

$$\sum_{k \ll ja} \|T_{jk}\| \leqslant \sum_{k=0}^{\infty} \sum_{r=0}^{\infty} \min\big(2^{-\frac{1}{2}k(1+\frac{1}{a})} 2^{-\frac{1}{2}\cdot\frac{r}{a}}, |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}kn}\big). \tag{4.7}$$

(Here we observe that the finite number of contributions coming possibly from the negative values of r, can be bounded by a multiple of the contributions say of $r=0$, and hence can be absorbed into the right-hand side of (4.7) up to a multiplicative constant, which we ignore.) The summation on the right-hand side of (4.7) is performed first with respect to the index appearing with opposite signs in the two estimates on the right-hand side, in this case k. It is convenient to state the outcome in a general form, as we shall use it often in the sequel. For large $|\lambda|$, and for any α, β positive constants, we have

$$\sum_{k=0}^{\infty} \min\big(2^{-\frac{1}{2}\alpha k} 2^{-\frac{1}{2}M}, |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}\beta k}\big) \sim |\lambda|^{-\frac{1}{2}\cdot\frac{\alpha}{\alpha+\beta}} 2^{-\frac{1}{2}\cdot\frac{\beta M}{\alpha+\beta}}. \tag{4.8}$$

In the case of (4.7), we find

$$\sum_{k \ll aj} \|T_{jk}\| \leqslant \sum_{r=0}^{\infty} |\lambda|^{-\frac{1}{2}\delta} 2^{-\frac{1}{2}\cdot\frac{nr}{1+a+an}} \sim |\lambda|^{-\frac{1}{2}\delta}$$

with δ the desired decay rate in (4.4).

In the range $k \gg aj$, the Hessian of the phase is of size

$$|S''| \sim 2^{-ajn}$$

on the support of the kernel of T_{jk}, and thus

$$\|T_{jk}\| \leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}ajn}. \tag{4.9}$$

Set now $k=aj+r$. Summing the estimates (4.5) and (4.9) in the same way as in the preceding case, we obtain

$$\begin{aligned}\sum_{k\gg aj} \|T_{jk}\| &\leqslant \sum_{j=0}^{\infty}\sum_{r=0}^{\infty} \min(2^{-\frac{1}{2}(1+a)j} 2^{-\frac{1}{2}r}, |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}ajn}) \\ &\sim \sum_{r=0}^{\infty} |\lambda|^{-\frac{1}{2}\delta} 2^{-\frac{1}{2}\cdot\frac{an}{1+a+an} r} \sim |\lambda|^{-\frac{1}{2}\delta}.\end{aligned} \tag{4.10}$$

Finally, we turn to the case $k \sim aj$. In this case, for each j, there is only a finitely bounded number of operators T_{jk}. Furthermore, for $|j-j'|$ sufficiently large, the projections of the supports of T_{jk} and $T_{j'k'}$ on both the x- and the y-axes, will be disjoint for (j,k) and (j',k') in this range. Thus the operators T_{jk} and $T_{j'k'}$ are orthogonal, and it suffices to establish the estimate

$$\|T_{jk}\| \leqslant |\lambda|^{-\frac{1}{2}\delta} \tag{4.11}$$

individually for each operator T_{jk}.

We introduce then a finer partition of the operator T_{jk},

$$\begin{aligned}T_{jk} &= \sum_m T_{jk}^m + \sum_m \widetilde{T}_{jk}^m, \\ T_{jk}^m f(x) &= \int_{-\infty}^{\infty} e^{i\lambda S(x,y)} \chi_k(y)\chi_j(x)\chi_m(y-x^a)\chi(x,y) f(y)\,dy, \\ \widetilde{T}_{jk}^m f(x) &= \int_{-\infty}^{\infty} e^{i\lambda S(x,y)} \chi_k(y)\chi_j(x)\chi_m(-(y-x^a))\chi(x,y) f(y)\,dy.\end{aligned} \tag{4.12}$$

We separate in this way the region where $y-x^a$ is positive from the region where it is negative. It suffices to consider T_{jk}^m and the positive region, as the other region can be treated in exactly the same way. Evidently T_{jk}^m is only different from 0 for $m \geqslant aj-C$, for some small integer C. In presence of the cut-off function $\chi_m(y-x^a)$, we have the following bounds on the cross sections Δx and Δy of the support of the kernel of T_{jk}^m:

$$\begin{aligned}\Delta y &\sim 2^{-m}, \\ \Delta x &\leqslant 2^{-m} 2^{-(1-a)j},\end{aligned} \tag{4.13}$$

where the second bound follows from differentiating the inequality $y-x^a \sim 2^{-m}$ with respect to x. The Operator van der Corput Lemma still applies to the operator T_{jk}^m, since the support of $\chi_m(y-x^a)$ is an admissible curved box, and both its y-cross section and the inverse of its derivatives with respect to y can be bounded uniformly by 2^{-m}. On the support of T_{jk}^m, we have

$$|S''| \sim 2^{-mn},$$

which leads to the size and oscillatory estimates for T_{jk}^m given by

$$\begin{aligned} \|T_{jk}^m\| &\leqslant 2^{-m} 2^{-\frac{1}{2}(1-a)j}, \\ \|T_{jk}^m\| &\leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}mn}. \end{aligned} \tag{4.14}$$

The constraint $m \geqslant aj$ is enforced as usual by setting

$$m = aj + M$$

so that the above estimates can be rewritten as

$$\begin{aligned} \|T_{jk}^m\| &\leqslant 2^{-\frac{1}{2}(1+a)j} 2^{-M}, \\ \|T_{jk}^m\| &\leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}ajn} 2^{\frac{1}{2}Mn}. \end{aligned} \tag{4.15}$$

To sum in M, we consider a convex means of these two estimates

$$\|T_{jk}^m\| \leqslant (|\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}ajn} 2^{\frac{1}{2}Mn})^{\theta} (2^{-\frac{1}{2}(1+a)j} 2^{-M})^{(1-\theta)} \tag{4.16}$$

for some θ between 0 and 1. Choose θ so as to cancel the j factors

$$an\theta = (1+a)(1-\theta). \tag{4.17}$$

We find that θ must coincide with the Newton decay rate δ of (4.4),

$$\delta = \frac{1+a}{1+a+an},$$

and the estimate (4.16) becomes

$$\|T_{jk}^m\| \leqslant |\lambda|^{-\frac{1}{2}\delta} 2^{-\left(1-\delta-\frac{1}{2}n\delta\right)M}. \tag{4.18}$$

To sum in M, we need to consider the sign of

$$1-\delta-\frac{1}{2}n\delta = \frac{1}{2}\cdot\frac{n(a-1)}{1+a+an}. \tag{4.19}$$

Thus the geometric series in M is rapidly decreasing and we can sum (4.18), obtaining the desired estimate (4.11), unless $a=1$.

This last case can be treated separately, directly from the original estimate (4.14),

$$\sum_{m\geqslant ja}\|T_{jk}^m\|\leqslant\sum_{m=0}^{\infty}\min(2^{-m},|\lambda|^{-\frac{1}{2}}2^{\frac{1}{2}mn})\leqslant|\lambda|^{-\frac{1}{2}\cdot\frac{2}{2+n}},$$

which is the correct estimate when $a=1$. The treatment of Model I is complete.

(b) *Model* II. This second model corresponds to the phase function

$$S''=(y-x^{a_1})\dots(y-x^{a_n}),$$

where the exponents a_i are distinct positive real numbers greater than or equal to 1, and we have ordered them as

$$1\leqslant a_1<a_2<\dots<a_n. \tag{4.20}$$

We begin by a preliminary discussion of the reduced Newton diagram of $S(x,y)$. Set

$$\begin{aligned}A_l&=a_1+\dots+a_l,\\ B_l&=n-l.\end{aligned} \tag{4.21}$$

Observation 1. The reduced Newton diagram of $S(x,y)$ has vertices at the points

$$(1+A_l,1+B_l),\quad 1\leqslant l\leqslant n, \tag{4.22}$$

and its boundary consists of, besides the two boundary lines parallel to the axes, the segments joining $(1+A_{l-1},1+B_{l-1})$ and $(1+A_l,1+B_l)$.

To see this, we expand the Hessian $S''(x,y)$ as

$$S''(x,y)=\sum_{l=0}^{n}(-1)^{n-l}\Big(\sum_{p_1<\dots<p_l}x^{a_{p_1}+\dots+a_{p_l}}\Big)y^{n-l}. \tag{4.23}$$

This shows that the points (A_l,B_l) are on the Newton diagram of $S''(x,y)$, and that any other index (m,n) occurring with non-vanishing coefficient in the Taylor expansion of $S''(x,y)$ already lies within at least one closed northeast quadrant with vertex at a point (A_l,B_l). In particular, the Newton diagram of $S''(x,y)$ is generated by these quadrants, while the reduced Newton diagram of $S(x,y)$ is generated by their translates by the vector $(1,1)$. Evidently these translates have corners at $(1+A_l,1+B_l)$.

To complete our claim, we shall show that each point $(1+A_{l+1},1+B_{l+1})$ lies strictly above the line joining $(1+A_{l-1},1+B_{l-1})$ and $(1+A_l,1+B_l)$, so that the region above the union of these segments is indeed a convex region. (More directly, the slope of the lth segment is $-a_l^{-1}$, and the a_l's are strictly increasing.)

The equation of the line joining $(1+A_{l-1}, 1+B_{l-1})$ and $(1+A_l, 1+B_l)$ is

$$q = -\frac{p}{a_l} + \frac{1+A_l+(1+B_l)a_l}{a_l}. \tag{4.24}$$

Substituting in the values $1+A_{l+1}$ for x and $1+B_{l+1}$ for y, we find that our ordering $a_l < a_{l+1}$ of the exponents a_l occurring in the factorization of $S''(x,y)$ is exactly the condition guaranteeing that $(1+A_{l+1}, 1+B_{l+1})$ lies strictly above the line of equation (4.24).

Observation 2. The Newton decay rate δ can be expressed as

$$\delta = \min_l \delta_l \tag{4.25}$$

with δ_l given by

$$\delta_l = \frac{1+a_l}{1+A_l+(1+B_l)a_l}. \tag{4.26}$$

This is an immediate consequence of the definition (1.2) for the Newton decay rate, since $(\delta_l^{-1}, \delta_l^{-1})$ is just the intersection of the line of equation (4.24) with the bisectrix $p=q$.

Observation 3. If the Newton decay rate δ is achieved by the segment joining $(1+A_{l-1}, 1+B_{l-1})$ and $(1+A_l, 1+B_l)$, i.e., if $\delta=\delta_l$, for some particular index l, then we must have for that index

$$B_l \leqslant A_l. \tag{4.27}$$

In fact, $\delta=\delta_l$ just means that the segment joining $(1+A_{l-1}, 1+B_{l-1})$ and $(1+A_l, 1+B_l)$ is a main boundary segment, that is, a boundary segment where the line $x=y$ meets the boundary of the reduced Newton diagram. Its right end point $(1+A_l, 1+B_l)$ must then be in the half-plane $y \leqslant x$, which translates into the inequality (4.27).

We now derive the desired estimate $\|T\| = O(|\lambda|^{-\frac{1}{2}\delta})$ for the operator T. As before, we decompose T into operators T_{jk}, and consider only the region $x>0, y>0$.

In the range $k \ll a_1 j$, the Hessian is of size $|S''| \sim 2^{-kn}$, and we have the two estimates

$$\|T_{jk}\| \leqslant 2^{-\frac{1}{2}(j+k)},$$
$$\|T_{jk}\| \leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}kn}.$$

Setting as before $[a_1 j] = k+r$, we obtain

$$\begin{aligned} \sum_{k \ll a_1 j} \|T_{jk}\| &\leqslant \sum_{k=0}^{\infty}\sum_{r=0}^{\infty} \min\left(2^{-\frac{1}{2}\cdot\left(1+\frac{1}{a_1}\right)k} 2^{-\frac{1}{2}\cdot\frac{r}{a_1}}, |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}kn}\right) \\ &\sim \sum_{r=0}^{\infty} |\lambda|^{-\frac{1}{2}\cdot\frac{1+a_1}{1+a_1+na_1}} 2^{-\frac{1}{2}\cdot\frac{nr}{1+a_1+na_1}} \\ &\sim |\lambda|^{-\frac{1}{2}\cdot\frac{1+a_1}{1+a_1+na_1}}. \end{aligned} \tag{4.28}$$

We recognize the exponent on the right-hand side as $\delta_1 \geqslant \delta$.

In the range $k \gg a_n j$, the size and oscillatory estimates for T_{jk} are given by

$$\|T_{jk}\| \leqslant 2^{-\frac{1}{2}(j+k)},$$
$$\|T_{jk}\| \leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}A_n j}.$$

Setting $k=[a_n j]+r$, the summation in j and k leads to

$$\begin{aligned}\sum_{k\gg a_n j} \|T_{jk}\| &\leqslant \sum_{j=0}^{\infty}\sum_{r=0}^{\infty} \min(2^{-\frac{1}{2}(1+a_n)j} 2^{-\frac{1}{2}r}, |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}A_n j}) \\ &\sim \sum_{r=0}^{\infty} |\lambda|^{-\frac{1}{2}\cdot\frac{1+a_n}{1+A_n+a_n}} 2^{-\frac{1}{2}\cdot\frac{A_n r}{1+A_n+a_n}} \\ &\sim |\lambda|^{-\frac{1}{2}\cdot\frac{1+a_n}{1+A_n+a_n}}\end{aligned} \tag{4.29}$$

with A_n defined by (4.21). The rate in (4.29) is exactly δ_n, which is again $\geqslant \delta$.

We consider now the range $a_l j \ll k \ll a_{l+1} j$. The factors $y-x^{a_i}$ are then of size $2^{-a_i j}$ or 2^{-k} respectively for $i \leqslant l$ and $i \geqslant l+1$, so that

$$|S''| \sim 2^{-A_l j} 2^{-B_l k}$$

with A_l, B_l defined as in (4.21). The size and oscillatory estimates for $\|T_{jk}\|$ are then

$$\begin{aligned}\|T_{jk}\| &\leqslant 2^{-\frac{1}{2}(j+k)}, \\ \|T_{jk}\| &\leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}A_l j} 2^{\frac{1}{2}B_l k}.\end{aligned} \tag{4.30}$$

To sum in the range $a_l j \ll k \ll a_{l+1} j$, we have to consider three different cases, depending on whether we have $B_l < A_l$, $A_l < B_l$ or $A_l = B_l$.

In the first case, where $B_l < A_l$, we set

$$k = [a_l j] + r, \quad r \geqslant 0,$$

in terms of which the preceding estimates become

$$\begin{aligned}\|T_{jk}\| &\leqslant 2^{-\frac{1}{2}(1+a_l)j} 2^{-\frac{1}{2}r}, \\ \|T_{jk}\| &\leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}(A_l+B_l a_l)j} 2^{\frac{1}{2}B_l r}.\end{aligned} \tag{4.31}$$

We shall take the convex means of these two estimates,

$$\|T_{jk}\| \leqslant [|\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}(A_l+B_l a_l)j} 2^{\frac{1}{2}B_l r}]^{\theta} [2^{-\frac{1}{2}(1+a_l)j} 2^{-\frac{1}{2}r}]^{(1-\theta)}, \tag{4.32}$$

which annihilates the j-factors. This determines θ as the solution of the equation

$$(A_l+B_l a_l)\theta=(1+a_l)(1-\theta). \tag{4.33}$$

Thus θ coincides exactly with δ_l,

$$\theta=\frac{1+a_l}{1+A_l+(B_l+1)a_l}=\delta_l. \tag{4.34}$$

The estimate (4.32) for $\|T_{jk}\|$ reduces to

$$\|T_{jk}\|\leqslant|\lambda|^{-\frac{1}{2}\delta_l}2^{-\frac{1}{2}(1-\delta_l-B_l\delta_l)r}. \tag{4.35}$$

For a fixed r, this same bound holds for the sum over j,k satisfying $k=[a_l j]+r$, since the operators T_{jk} and $T_{j'k'}$ have disjoint x- and y-supports, and hence are orthogonal when $|j-j'|$ is larger than some fixed constant

$$\Big\|\sum_{k=[a_l j]+r} T_{jk}\Big\|\leqslant|\lambda|^{-\frac{1}{2}\delta_l}2^{-\frac{1}{2}(1-\delta_l-B_l\delta_l)r}. \tag{4.36}$$

In the case we are considering, $B_l<A_l$, and the geometric series in r in (4.36) is convergent, since

$$1-\delta_l-B_l\delta_l=1-\frac{(1+B_l)(1+a_l)}{1+A_l+(1+B_l)a_l}=\frac{A_l-B_l}{1+A_l+(1+B_l)a_l}>0. \tag{4.37}$$

We obtain in this way the desired estimate

$$\sum_{r=0}^{\infty}\Big\|\sum_{k=[a_l j]+r} T_{jk}\Big\|\leqslant|\lambda|^{-\frac{1}{2}\delta_l}\leqslant|\lambda|^{-\frac{1}{2}\delta}.$$

Consider next the case $A_l<B_l$. We set instead

$$j=\left[\frac{1}{a_{l+1}}k\right]+r,\quad r\geqslant 0,$$

and rewrite the two basic estimates (4.30) as

$$\begin{aligned}\|T_{jk}\|&\leqslant 2^{-\frac{1}{2}\cdot\left(1+\frac{1}{a_{l+1}}\right)k}2^{-\frac{1}{2}r},\\ \|T_{jk}\|&\leqslant|\lambda|^{-\frac{1}{2}}2^{\frac{1}{2}\cdot\left(\frac{A_l}{a_{l+1}}+B_l\right)k}2^{\frac{1}{2}A_l r}.\end{aligned} \tag{4.38}$$

The convex combination θ annihilating the k-factors is given by

$$\left(\frac{A_l}{a_{l+1}}+B_l\right)\theta=\left(1+\frac{1}{a_{l+1}}\right)(1-\theta).$$

We find this time

$$\theta=\frac{1+a_{l+1}}{1+A_l+(B_l+1)a_{l+1}}=\frac{1+a_{l+1}}{1+A_{l+1}+(B_{l+1}+1)a_{l+1}}=\delta_{l+1}. \tag{4.39}$$

Again, by orthogonality, the sum of the T_{jk}'s over j, k satisfying $j=[k/a_{l+1}]+r$ for fixed r satisfies the same estimate as the individual T_{jk}'s,

$$\Big\|\sum_{j=[k/a_{l+1}]+r} T_{jk}\Big\|\leqslant|\lambda|^{-\frac{1}{2}\delta_{l+1}}2^{-\frac{1}{2}(1-\delta_{l+1}-A_l\delta_{l+1})r}. \tag{4.40}$$

In the range $A_l<B_l$, the above geometric series in r is again summable, since

$$1-\delta_{l+1}-A_l\delta_{l+1}=1-\frac{(1+a_{l+1})(1+A_l)}{1+A_l+(B_l+1)a_{l+1}}=a_{l+1}\frac{B_l-A_l}{1+A_l+(B_l+1)a_{l+1}}>0. \tag{4.41}$$

The desired estimate follows,

$$\sum_{r=0}^{\infty}\Big\|\sum_{j=[k/a_{l+1}]+r} T_{jk}\Big\|\leqslant|\lambda|^{-\frac{1}{2}\delta_{l+1}}\leqslant|\lambda|^{-\frac{1}{2}\delta}.$$

The third subcase in the range $a_l j\ll k\ll a_{l+1}j$ that we need to consider is the case when $A_l=B_l$. In this case, we simply observe that the basic size and oscillatory estimates of (4.30) can be rewritten as

$$\begin{aligned}\|T_{jk}\|&\leqslant 2^{-\frac{1}{2}(j+k)},\\ \|T_{jk}\|&\leqslant|\lambda|^{-\frac{1}{2}}2^{\frac{1}{2}A_l(j+k)}.\end{aligned} \tag{4.42}$$

The sole dependence on $j+k$ of both estimates in (4.42) allows us to sum diagonally over all j, k satisfying $j+k=i$ for fixed i. Indeed, along such diagonals, $|j-j'|\gg 1$ implies that $|k-k'|\gg 1$, and thus T_{jk} and $T_{j'k'}$ are orthogonal for $|j-j'|$ large. By orthogonality, the sum over j, k satisfying $j+k=i$ admits then the same bound as each individual summand. Thus

$$\Big\|\sum_{j+k=i} T_{jk}\Big\|\leqslant\min(2^{-\frac{1}{2}i},|\lambda|^{-\frac{1}{2}}2^{\frac{1}{2}A_l i}). \tag{4.43}$$

This leads immediately to the following bound for the sum in i:

$$\sum_{i=0}^{\infty}\Big\|\sum_{j+k=i} T_{jk}\Big\|\leqslant|\lambda|^{-\frac{1}{2}\cdot\frac{1}{1+A_l}}. \tag{4.44}$$

Since in the case $A_l=B_l$, we may write

$$\delta_l=\frac{1+a_l}{1+A_l+(B_l+1)a_l}=\frac{1+a_l}{(1+A_l)(1+a_l)}=\frac{1}{1+A_l},$$

the preceding bound is again of the form $O(|\lambda|^{-\frac{1}{2}\delta_l})$, and our treatment of the range $a_l j \ll k \ll a_{l+1} j$ is complete.

The only range which remains to be treated is the range $k \sim a_l j$. In this range, we begin by observing that by suitably restricting the neighborhood of the origin, we may assume that both k and j are so large that

$$|j(a_{l+1}-a_l)| > \max_l a_l + 10$$

and thus the ranges of the form $k \sim a_l j$ and $k \sim a_{l+1} j$ are well separated. As before, the finite number of k's in this range and the orthogonality of T_{jk} and $T_{j'k'}$ for $|j-j'|$ large (this implies that $|k-k'|$ is large as well) allow us to restrict ourselves to the proof of the estimate $\|T_{jk}\| \leqslant |\lambda|^{\frac{1}{2}\delta}$ for individual (j,k)'s.

Again, as in the third range considered in Model I, we introduce a further decomposition of the form (4.12) with $\sum_m \sum_{\sigma=\pm} \chi_m(\sigma(y-x^{a_l}))$, and study the bounds for the resulting T_{jk}^m. Only the range $m \geqslant a_l j$ contributes.

The analogues of the estimates (4.14) are in the present case

$$\begin{aligned} \|T_{jk}^m\| &\leqslant 2^{-m} 2^{-\frac{1}{2}(1-a_l)j}, \\ \|T_{jk}^m\| &\leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}A_{l-1}j} 2^{\frac{1}{2}m} 2^{\frac{1}{2}B_l a_l j}. \end{aligned} \tag{4.45}$$

Set $m=[a_l j]+M$. As explained before, we may consider only M non-negative integer, up to a suitable multiplicative constant in the subsequent estimates. In terms of M, (4.45) reduces to the following analogue of (4.15):

$$\begin{aligned} \|T_{jk}^m\| &\leqslant 2^{-\frac{1}{2}(1+a_l)j} 2^{-M}, \\ \|T_{jk}^m\| &\leqslant \lambda^{-\frac{1}{2}} 2^{\frac{1}{2}(A_l+B_l a_l)j} 2^{\frac{1}{2}M}. \end{aligned} \tag{4.46}$$

The convex combination θ annihilating the j-factors is given by

$$(A_l+B_l a_l)\theta = (1+a_l)(1-\theta)$$

and thus

$$\theta = \delta_l$$

as before. The resulting estimate for T_{jk}^m is

$$\|T_{jk}^m\| \leqslant |\lambda|^{-\frac{1}{2}\delta_l} 2^{-\frac{1}{2}(2-3\delta_l)M}. \tag{4.47}$$

We need to check the sign of $2-3\delta_l$. This expression can be rewritten as

$$2-3\delta_l = \frac{2A_{l-1}+(2B_l+1)a_l-1}{1+A_l+(1+B_l)a_l}. \tag{4.48}$$

Since a_l is assumed to be $\geqslant 1$, the right-hand side is always strictly positive unless $a_l=1$ and $B_l=0$, which means that we have only a single factor in the factorization for $S''(x,y)$, of the form $S''(x,y)=y-x$. This last case has already been treated in Model I, with the desired decay $\|T\|\leqslant|\lambda|^{-1/3}$. In the remaining cases, the geometric series in M is convergent, and we do obtain

$$\sum_{m\geqslant a_l j}^{\infty} \|T_{jk}^m\| \leqslant |\lambda|^{-\frac{1}{2}\delta_l} \leqslant |\lambda|^{-\frac{1}{2}\delta}.$$

The treatment of Model II is complete.

(c) *Model* III. In Model III, we consider the simplest case of phase functions which may have the same leading exponents, but with different coefficients,

$$S'' = (y-c_1x^a)\dots(y-c_Nx^a), \quad c_j\neq c_k \text{ for } j\neq k.$$

We begin by observing that for

$$k < aj - K(c_1, \dots, c_N; a)$$

(which we again denote by $k\ll aj$), there is no major cancellation in any of the factors $y-c_ix^a$. Thus we have

$$|S''| \sim 2^{-kN},$$

and the same summation techniques used in the case $k\ll a_1 j$ of Model II apply to yield the desired estimate

$$\sum_{k\ll aj} \|T_{jk}\| \leqslant |\lambda|^{-\frac{1}{2}\cdot\frac{1+a}{1+a+aN}}.$$

Similarly, the range $k>aj+\widetilde{K}(c_1,\dots,c_N;a)$ (abbreviated by $k\gg aj$) is treated along the lines of the case $k\gg a_n j$ of Model II, and presents no new difficulty.

We consider now the range $-K(c_1,\dots,c_n;a)+aj\leqslant k\leqslant \widetilde{K}(c_1,\dots,c_N;a)+aj$, which we denote by $k\sim aj$. By orthogonality, it suffices to show that $\|T_{jk}\|\leqslant|\lambda|^{-\frac{1}{2}\cdot\frac{1+a}{1+a+aN}}$ for each (j,k). In this case, we need a further decomposition. Select any of the factors in S'', say, $y-c_1x^a$, and introduce a partition of unity $\sum_m \chi_m(y-c_1x^a)$, leading to an analogue of (4.12). Thus, in the support of the kernel of T_{jk}^m, we have

$$y-c_1x^a\sim 2^{-m}, \quad m\geqslant aj.$$

For a suitably large $K_1(c_1,\dots,c_N;a)$, the condition

$$m > aj + K_1(c_1, \dots, c_N; a)$$

implies that for $i\neq 1$,

$$|y-c_ix^a|=|y-c_1x^a+(c_1-c_i)x^a|\sim 2^{-aj}.$$

Thus the bounds $|S''|\sim 2^{-(N-1)aj}2^{-m}$ are sharp, and the familiar two estimates for T_{jk}^m are in this case

$$\begin{aligned}\|T_{jk}^m\| &\leqslant 2^{-m}2^{-\frac{1}{2}(1-a)j},\\ \|T_{jk}^m\| &\leqslant |\lambda|^{-\frac{1}{2}}2^{\frac{1}{2}m}2^{\frac{1}{2}(N-1)aj}.\end{aligned}\tag{4.49}$$

This is the same estimate as in (4.45), say with $a_1=...=a_l=a$, $l=N$, and m replaced by M. Thus it implies the desired decay rate $|\lambda|^{-\frac{1}{2}\delta}$. We may assume then that

$$m\sim aj \quad\text{and}\quad y-c_1x^a\sim 2^{-m}.$$

In particular, the number of such m's is boundedly finite, and

$$|S''|\sim 2^{-aj}|y-c_2x^a|\,...\,|y-c_Nx^a|.\tag{4.50}$$

Thus we are reduced to the same case, with one factor less. If we go on in this manner, we can keep eliminating more factors in S''. Each step requires the insertion of a further cut-off $\chi_{\tilde{m}}(y-c_ix^a)$, which does not affect the applicability of the Operator van der Corput Lemma. The estimates we encounter are of the form (4.49), and can be handled just as in that case. Finally, we arrive at the case when all factors $y-c_ix^a$ have been peeled off, and $|S''|\sim 2^{-Naj}$, so that $\|T_{ik}^{m_1,\dots,m_N}\|\leqslant|\lambda|^{-\frac{1}{2}}2^{\frac{1}{2}Naj}$, with $m_i\sim aj$ denoting the decompositions inherent to each step. On the support side, we have $\Delta x\leqslant 2^{-m}2^{-(1-a)j}$, $\Delta y\leqslant 2^{-m}$, which implies that $\|T_{jk}^{m_1,\dots,m_N}\|\leqslant 2^{-m}2^{\frac{1}{2}(1-a)j}\sim 2^{-\frac{1}{2}(1+a)j}$. Altogether,

$$\|T_{jk}^{m_1,\dots,m_N}\|\leqslant\min(2^{-\frac{1}{2}(1+a)j},|\lambda|^{-\frac{1}{2}}2^{\frac{1}{2}Naj})\sim|\lambda|^{-\frac{1}{2}\cdot\frac{1+a}{1+a+aN}},$$

which is the desired estimate.

(d) *Model* IV. We come now to the basic situation where several roots of the Hessian S'' can be highly tangent, while remaining nevertheless distinct, i.e.,

$$S''=(y-x^a-r_1(x)x^{b_1})\,...\,(y-x^a-r_N(x)x^{b_N})$$

with $N\geqslant 2$, $1\leqslant a<b_1<b_2<...<b_N$, and $r_1(x),...,r_N(x)$ functions not vanishing at 0. It is not difficult to verify that the ranges $k\ll aj$ and $k\gg aj$ of our basic decompositions lead as before to the correct gain $\sum\|T_{jk}\|\leqslant|\lambda|^{-\frac{1}{2}\delta}$, with $\delta=(1+a)/(1+a+aN)$. Thus we need consider only the range $k\sim aj$, where it suffices to establish the desired estimate for each individual T_{jk}.

We decompose T_{jk} into T_{jk}^m by introducing a partition of unity $\chi_m(y-x^a-r_N(x)x^{b_N})$ (restricting ourselves again to discuss only the side where $y-x^a-r_N(x)x^{b_N}>0$). Thus we have

$$y-x^a-r_N(x)x^{b_N}\sim 2^{-m}, \quad m\geqslant aj,$$

on the support of the kernel of T_{jk}^m. If $m\gg b_{N-1}j$, we have for each of the factors

$$|y-x^a-r_i(x)x^{b_i}|=|(y-x^a-r_N(x)x^{b_N})+(r_N(x)x^{b_N}-r_i(x)x^{b_i})|\sim 2^{-jb_i}$$

for $i\leqslant N-1$. The size of $|S''|$ is then

$$|S''|\sim 2^{-(b_1+\ldots+b_{N-1})j}2^{-m}.$$

By restricting the support of T to be small enough, we may insure that the factors $y-x^a-r_i(x)x^{b_i}$ are all monotone functions of x. Thus the Operator van der Corput Lemma is applicable, and we have the two estimates

$$\begin{aligned}\|T_{jk}^m\| &\leqslant 2^{-m}2^{-\frac{1}{2}(1-a)j},\\ \|T_{jk}^m\| &\leqslant |\lambda|^{-\frac{1}{2}}2^{\frac{1}{2}(b_1+\ldots+b_{N-1})j}2^{\frac{1}{2}m},\end{aligned} \tag{4.51}$$

with the first estimate a consequence of familiar estimates for the cross sections of the support of T_{jk}^m.

To sum in m, we note that (4.51) is the same type of estimates as in e.g. (4.45). We summarize such estimates in the following lemma.

LEMMA 3. *Let the terms $\mathcal{T}_m$ of a series be bounded by*

$$\begin{aligned}\mathcal{T}_m &\leqslant 2^{-m}2^{-\frac{1}{2}(1-a)j},\\ \mathcal{T}_m &\leqslant |\lambda|^{-\frac{1}{2}}2^{\frac{1}{2}Kj}2^{\frac{1}{2}\varrho m},\end{aligned}$$

for $m\geqslant bj$, with a constant $b\geqslant a$, and non-negative constants K and ϱ. Assume that $a>1$, or that $K>0$ if $a=1$. Then

$$\sum_{m\geqslant bj}\mathcal{T}_m\leqslant |\lambda|^{-\frac{1}{2}\cdot\frac{1+2b-a}{1+K+(\varrho+2)b-a}}.$$

Proof of Lemma 3. In terms of r, $m=bj+r$, the estimates in the hypothesis of Lemma 3 can be expressed as

$$\begin{aligned}\mathcal{T}^m &\leqslant 2^{-\frac{1}{2}(1+2b-a)j}2^{-r},\\ \mathcal{T}^m &\leqslant |\lambda|^{-\frac{1}{2}}2^{\frac{1}{2}(K+\varrho b)j}2^{\frac{1}{2}\varrho r}.\end{aligned}$$

The convex combination θ annihilating the j-factors is

$$\theta = \frac{1+2b-a}{1+K+(\varrho+2)b-a}.$$

We need to check that the resulting estimate can be summed in r. This is the case when $-\frac{1}{2}\varrho\theta+(1-\theta)>0$, which works out to be

$$K+\tfrac{1}{2}\varrho(a-1)>0.$$

This condition is satisfied under the hypotheses of Lemma 3. The proof of Lemma 3 is complete.

Returning now to the estimates (4.51) for $\|T^m_{jk}\|$, we can apply Lemma 3 and obtain

$$\sum_{m\geqslant b_{N-1}j}\|T^m_{jk}\|\leqslant|\lambda|^{-\frac{1}{2}\cdot\frac{1+2b_{N-1}-a}{1+b_1+\ldots+b_{N-1}+3b_{N-1}-a}}. \tag{4.52}$$

To compare this gain to the desired gain δ, we need the following lemma, which is a key tool for the case of highly tangent roots:

LEMMA 4. *Let $1\leqslant a\leqslant b_1\leqslant b_2\leqslant\ldots b_{N-1}\leqslant b_N$ be a finite sequence of numbers, and set*

$$\delta_p(b)=\frac{1+2b-a}{1+b_1+\ldots+b_{p-1}+(N-p+3)b-a}.$$

Then

$$\delta_N(b_N)\geqslant\delta_{N-1}(b_{N-1})\geqslant\ldots\geqslant\delta_1(b_1)\geqslant\delta_1(a)=\delta, \tag{4.53}$$

with strict inequality

$$\delta_p(b_p)>\delta_{p-1}(b_{p-1})\quad\text{when } b_p>b_{p-1}\text{ and } p\geqslant 2. \tag{4.54}$$

When $a>1$, we also have the strict inequality

$$\delta_1(b_1)>\delta_1(a)\quad\text{when } b_1>a, \tag{4.55}$$

while when $a=1$, we have identically for all b_1

$$\delta_1(b_1)=\frac{2}{N+2}. \tag{4.56}$$

Proof of Lemma 4. In fact, the derivative of $\delta_p(b)$ viewed as a function of b is proportional to

$$2(b_1+\ldots+b_{p-1})+(N-p+1)(a-1), \tag{4.57}$$

up to a strictly positive factor. Thus $\delta_p(b)$ is an increasing function of b, and since $\delta_p(b_p) \geqslant \delta_p(b_{p-1}) = \delta_{p-1}(b_{p-1})$, we can iterate and obtain all the inequalities listed in (4.53), ending with $\delta_1(b_1) \geqslant \delta_1(a) = \delta$. Evidently, the expression (4.57) is strictly positive when $p \geqslant 2$, or when $a > 1$, which accounts for (4.54) and (4.55). Finally, when $p=1$ and $a=1$, a direct calculation gives (4.56). Lemma 4 is proved.

With Lemma 4, it is clear that the gain we derived in (4.52) is at least as good as the desired gain.

Next, the range $jb_p \ll m \ll jb_{p+1}$ requires only a few modifications, so we shall be succinct. In this range, $|S''| \sim 2^{-(b_1+\ldots+b_p)j} 2^{-m(N-p)}$, and we just need to sum the two estimates

$$\begin{aligned} \|T_{jk}^m\| &\leqslant 2^{-m} 2^{-\frac{1}{2}(1-a)j}, \\ \|T_{jk}^m\| &\leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}(b_1+\ldots+b_p)j} 2^{\frac{1}{2}(N-p)m}. \end{aligned} \tag{4.58}$$

Setting $m = [jb_p] + r$, $r = jM$, we consider a θ_M-convex combination of the estimates in (4.58),

$$\|T_{jk}^m\| \leqslant [|\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}(b_1+\ldots+b_p+(N-p)(b_p+M))j}]^{\theta_M} [2^{-\frac{1}{2}(1+2b_p+2M-a)j}]^{(1-\theta_M)}. \tag{4.59}$$

Let θ_M^* be defined so as to annihilate the j-factors. It is given by

$$\theta_M^* = \frac{1+2(b_p+M)-a}{1+b_1+\ldots+b_p+(N-p+2)(b_p+M)-a}.$$

As in the proof of Lemma 4, this expression is an increasing function of M. Thus

$$\theta_M^* \geqslant \theta_0^* = \frac{1+2b_p-a}{1+b_1+\ldots+b_p+(N-p+2)b_p-a} = \delta_p(b_p).$$

We can now quote directly Lemma 4, and conclude that $\theta_0^* > \delta$ unless $a=1$ and $p=1$. Postponing this particular case for the moment, we may choose θ_M to be $\theta_M = \theta_M^* - \varepsilon > \delta_p(b_p)$ for a small positive ε, and write

$$\begin{aligned} \|T_{jk}^m\| &\leqslant |\lambda|^{-\frac{1}{2}\theta_M} 2^{-\frac{1}{2}(1+b_1+\ldots+b_p+(N-p+2)(b_p+M)-a)\varepsilon j} \\ &\sim |\lambda|^{-\frac{1}{2}\theta_M} 2^{-\frac{1}{2}(1+b_1+\ldots+b_p+(N-p+2)b_p-a)\varepsilon j} 2^{-\frac{1}{2}(N-p+2)\varepsilon r} \\ &\leqslant |\lambda|^{-\frac{1}{2}\delta_p(b_p)} 2^{-\frac{1}{2}(1+b_1+\ldots+b_p+(N-p+2)b_p-a)\varepsilon j} 2^{-\frac{1}{2}(N-p+2)\varepsilon r}. \end{aligned} \tag{4.60}$$

The series in r is a convergent geometric series, and we can conclude that

$$\sum_{m \gg jb_p} \|T_{jk}^m\| \leqslant |\lambda|^{-\frac{1}{2}\delta_p(b_p)} \leqslant |\lambda|^{-\frac{1}{2}\delta}.$$

The other end range, $aj \ll m \ll jb_1$ is even simpler. There $|S''| \sim 2^{-Nm}$, and thus

$$\|T^m_{jk}\| \leqslant 2^{-m} 2^{-\frac{1}{2}(1-a)j},$$
$$\|T^m_{jk}\| \leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}Nm}.$$

By Lemma 3, this leads to $\sum_{m \geqslant aj} \|T^m_{jk}\| \leqslant |\lambda|^{-\frac{1}{2} \cdot \frac{1+a}{1+a+Na}} = |\lambda|^{-\frac{1}{2}\delta}$.

We turn now to the special case where $a=1$ and $p=1$. Set $m=[b_1 j]+r$. The size and oscillatory estimates for T^m_{jk} reduce to

$$\begin{aligned} &\|T^m_{jk}\| \leqslant 2^{-m} = 2^{-b_1 j} 2^{-r}, \\ &\|T^m_{jk}\| \leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2} b_1 j} 2^{\frac{1}{2}(N-1)m} = |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2} N b_1 j} 2^{\frac{1}{2}(N-1)r}. \end{aligned} \tag{4.61}$$

The convex combination θ annihilating the j factors is

$$\tfrac{1}{2} N\theta = 1-\theta \quad \Leftrightarrow \quad \theta = \frac{2}{2+N} = \delta.$$

This leads to the desired estimate

$$\sum_{r=0}^{\infty} \|T^m_{jk}\| \leqslant \sum_{r=0}^{\infty} |\lambda|^{-\frac{1}{2}\delta} 2^{-\frac{1}{2+N} r} \sim |\lambda|^{-\frac{1}{2}\delta}.$$

Finally, we are left with considering the range $m \sim jb_p$, where the number of m's is boundedly finite. The size of the Hessian is controlled by

$$|S''| \sim 2^{-(N-p)b_p j} |y - x^a - r_1(x) x^{b_1}| \dots |y - x^a - r_p(x) x^{b_p}|,$$

which is of the same form as the original S'', but with fewer factors. As in Model III, continuing this peeling off process leads ultimately to summands where S'' is of constant size. We can then verify as before that the resulting decay rate is greater than or equal to δ.

(e) *Model* V. We consider now a Hessian of the form

$$S'' = \prod_{\alpha}^{n} \prod_{p=1}^{N_\alpha} (y - c_\alpha x^a - r_{\alpha p}(x) x^{b_{\alpha p}}),$$

with $0 \neq c_j \neq c_k$ if $j \neq k$, $r_{\alpha p}(0) \neq 0$. We also assume that for each fixed α, the exponents $b_{\alpha p}$ are all distinct. In this case, the gain sought is $\delta = (1+a)/(1+a+aN)$, where we have set $N = N_1 + \dots + N_n$. By convention, we order the $b_{\alpha p}$'s so that

$$b_{\alpha 1} < b_{\alpha 2} < \dots < b_{\alpha N_\alpha}.$$

As usual, the range $k\sim aj$ is the only one requiring a careful discussion. In this range, we choose an α, say $\alpha=1$, and introduce a further decomposition in $y-c_1x^a-r_{1N_1}x^{b_{1N_1}}$, so that, once again on the support of the kernel of each of the components T_{jk}^m of T_{jk} ($m\geqslant aj$), we have

$$y-c_1x^a-r_{1N_1}(x)x^{b_{1N_1}}\sim 2^{-m}.$$

If $m\gg aj+K[c_\alpha, b_{\alpha p}; r_{\alpha p}]$, then

$$|S''|\sim 2^{-(N_1+\ldots+N_n)aj}|y-c_1x^a-r_{11}(x)x^{b_{11}}|\ldots|y-c_1x^a-r_{1N_1}(x)x^{b_{1N_1}}|. \tag{4.62}$$

This is of the form of Model IV, only with the additional prefactor $2^{-(N_1+\ldots+N_n)aj}$. This factor is easily incorporated in the arguments for Model IV. Without it, the gain $\tilde{\delta}$ for the phase function (4.62) would be

$$\frac{1+a}{1+a+aN_1}.$$

With a factor 2^{-Kj} in front of S'', the gain gets shifted to (cf. Lemma 3)

$$\frac{1+a}{1+K+a+aN_1}.$$

In the present case we have $K=a(N_2+\ldots+N_n)$, and the preceding formula does become $(1+a)/(1+a+aN)=\delta$.

We are reduced then to the case $m\leqslant aj+K[c_\alpha, b_{\alpha p}; r_{\alpha p}]$, and in particular $m\sim aj$. This implies that the number of m's is finite, and thus we can fix m. Now for each q,

$$|y-c_1x^a-r_{1q}(x)x^{b_{1q}}|=|(y-c_1x^a-r_{1N_1}(x)x^{b_{1N_1}})+(r_{1N_1}(x)x^{b_{1N_1}}-r_{1q}(x)x^{b_{1q}})|\sim 2^{-aj}.$$

Thus we have the following estimate for the Hessian:

$$|S''|\sim 2^{-N_1aj}\left|\prod_{\alpha=2}^{n}\prod_{p=1}^{N_\alpha}(y-c_\alpha x^a-r_{\alpha p}(x)x^{b_{\alpha p}})\right|. \tag{4.63}$$

This is of the same form as the original estimate, but with one less factor in α. We can now proceed inductively. For example, decomposing $y-c_2x^a-r_{2N_2}x^{b_{2N_2}}\sim 2^{-m_2}$, for $m_2\gg aj$, we are led to Model IV with

$$|S''|\sim 2^{-(N-N_2)aj}\left|\prod_{p=1}^{N_2}(y-c_2x^a-r_{2p}(x)x^{b_{2p}})\right|.$$

The correct estimate follows again from Lemma 3. On the other hand, in the range $m_2\sim aj$, we reduce S'' by yet another factor α. Ultimately, we arrive at

$$|S''|\sim 2^{-Naj}$$

and the two estimates (all the m_j's are $\sim aj$)

$$\|T_{jk}^{m_1,\dots,m_n}\| \leqslant 2^{-\frac{1}{2}(1+a)j},$$
$$\|T_{jk}^{m_1,\dots,m_n}\| \leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}Naj}.$$

In particular,

$$\|T_{jk}^{m_1,\dots,m_n}\| \leqslant \min(2^{-\frac{1}{2}(1+a)j}, |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2}Naj}) \sim |\lambda|^{-\frac{1}{2}\cdot\frac{1+a}{1+a+aN}}.$$

Our discussion of Model V is complete.

(f) *Complex roots.* We shall now indicate how the above arguments are modified in the presence of complex roots. Let a factor of S'' be given by $y-r(x)$, and let $r(x)=cx^a+$ higher powers. There are then two possibilities:

If $c\in\mathbf{R}$, then

$$|y-r(x)| \sim |y-\operatorname{Re} r(x)|+|\operatorname{Im} r(x)| \sim |y-\operatorname{Re} r(x)|+|x|^b \tag{4.64}$$

for some $b>a$. We note that the function $\operatorname{Re} r(x)$ is a Puiseux series in x with real coefficients, and that its leading exponent is a, the same as the leading exponent of $r(x)$.

Otherwise $c\notin\mathbf{R}$. Then we have

$$\begin{aligned}|y-r(x)| &\sim |y-\operatorname{Re} r(x)|+|\operatorname{Im} r(x)| \sim \varepsilon|y-\operatorname{Re} r(x)|+|x|^a\\ &\geqslant \varepsilon||y|-C|x|^a|+|x|^a \geqslant \varepsilon(|y|+|x|^a),\end{aligned} \tag{4.65}$$

where ε is a generic notation for a constant which can be fixed at an arbitrarily small value. Since the reverse inequality $|y-r(x)|\leqslant C(|y|+|x|^a)$ is evident, we have in this case

$$|y-r(x)| \sim |y|+|x|^a. \tag{4.66}$$

We can now see easily how the arguments for the boundedness of T apply essentially without any change in presence of complex roots. Consider e.g. Model II, with any factor $|y-x^{a_l}|$ possibly replaced by either $|y|+|x|^{a_l}$ or by $|y-x^{a_l}|+|x|^{b_l}$ for some $b_l>a_l$. In the first case, the argument is even simplified, since the bounds for T_{jk} are unaffected in the range $a_l j\ll k\ll a_{l+1}j$. Even in the range $k\sim a_l j$, there can be no cancellation between the terms $|y|$ and $|x|^{a_l}$, so that this range does not even require a separate treatment. Alternatively, if we set as before $|y|+|x|^{a_l}\sim 2^{-m}$ in the range $k\sim a_l j$, the exponent m satisfies $m\sim a_l j$, and hence can take on only a finite number of values.

In the second case, the factor $|y-x^{a_l}|$ has been replaced by the better bound $|y-x^{a_l}|+|x|^{b_l}$, e.g.,

$$|S''| \sim (|y-x^{a_l}|+|x|^{b_l})\prod_{\mu\neq l}|y-x^{a_\mu}| \geqslant \prod_{\mu=1}^{l}|y-x^{a_\mu}|. \tag{4.67}$$

If we can apply the Operator van der Corput Lemma with only bounds from below for the Hessian of the phase, the desired estimates for all the T_{jk}'s, and hence for T, would follow at once. The Operator van der Corput Lemma requires however that the Hessian of the phase be uniformly bounded from *both* above and below, up to a multiplicative constant, by the same bound. This can be taken care of by the following minor variation of our argument. In the ranges $a_l j \ll k$ and $k \gg a_l j$, the bounds from above and below for S'' are unaffected by the presence of $|x|^{b_l}$, since there is no cancellation between y and x^{a_l}, and $|x|^{a_l}$ is much larger than $|x|^{b_l}$. Thus we need only consider the range $k \sim a_l j$. Set as before $y - x^{a_l} \sim 2^{-m}$, $m \geqslant a_l j$. Then the factor $|y - x^{a_l}| + |x|^{b_l}$ in the factorization of S'' admits the following bound from both above and below:

$$|y - x^{a_l}| + |x|^{b_l} \sim 2^{-\min(b_l j, m)}. \tag{4.68}$$

The Operator van der Corput Lemma can be applied, and leads to the same bound for $\|T_{jk}^m\|$ as in (4.45), with the contribution of $|y - x^{a_l}|$ being replaced by the contribution (4.68) of $|y - x^{a_l}| + |x|^{b_l}$. With the same notation as in (4.45), we have then

$$\|T_{jk}^m\| \leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2} A_{l-1} j} 2^{\frac{1}{2} \min(b_l j, m)} 2^{\frac{1}{2} B_l a_l j} \leqslant |\lambda|^{-\frac{1}{2}} 2^{\frac{1}{2} A_{l-1} j} 2^{\frac{1}{2} m} 2^{\frac{1}{2} B_l a_l j}. \tag{4.69}$$

The earlier proof now takes over without change, producing the same final bound for $\|T\|$, in terms of the exponents a_l.

5. Clusters of roots

(a) *Clusters of roots.* In the previous section, we had analyzed model cases where the roots of the Hessian can be classified by their leading exponent x^{a_l}. For each exponent a_l, we can have however a cluster of roots of the form $c^\alpha x^{a_l} + r_{lq}^\alpha x^{b_{lq}^\alpha}$. In the model cases, we had essentially assumed that all the coefficients $r_{lq}^\alpha(0)$ and the exponents b_{lq}^α were distinct, but this may very well not be the case in general. Rather, each term $r_{lq}^\alpha(0) x^{b_{lq}^\alpha}$ can be in turn the next leading coefficient for a smaller cluster of roots, which can be thought of as visible only with this finer resolution. Clearly, this process can repeat itself, until we reach, after a finite number of steps, a stage where all clusters consist each of a single root only, counted with its multiplicity. In order to deal more systematically with this picture, it is convenient to introduce the following notions.

Let all the distinct roots of S'', different from $x \equiv 0$ and $y \equiv 0$, be expanded as

$$r(x) = c_{l_1}^{\alpha_1} x^{a_{l_1}} + c_{l_1 l_2}^{\alpha_1 \alpha_2} x^{a_{l_1 l_2}^{\alpha_1}} + \ldots + c_{l_1 \ldots l_p}^{\alpha_1 \ldots \alpha_p} x^{a_{l_1 \ldots l_p}^{\alpha_1 \ldots \alpha_{p-1}}} + \ldots,$$

where

$$c_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_{p-1}\beta} \neq c_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_{p-1}\gamma} \quad \text{for } \beta\neq\gamma,$$
$$a_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_{p-1}} > a_{l_1\ldots l_{p-1}}^{\alpha_1\ldots\alpha_{p-2}},$$

and we have kept enough terms to distinguish between all the non-identical roots of S''.

By the cluster

$$\begin{bmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{bmatrix},$$

we shall designate all the roots $r(x)$, counted with their multiplicities, which satisfy

$$r(x)-(c_{l_1}^{\alpha_1}x^{a_{l_1}}+c_{l_1l_2}^{\alpha_1\alpha_2}x^{a_{l_1l_2}^{\alpha_1}}+\ldots+c_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_p}x^{a_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_{p-1}}})=O(x^b) \tag{5.1}$$

for some exponent $b>a_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_{p-1}}$. We also introduce the clusters

$$\begin{bmatrix} \alpha_1 & \ldots & \alpha_p & \cdot \\ l_1 & \ldots & l_{p-1} & l_p \end{bmatrix}$$

by

$$\begin{bmatrix} \alpha_1 & \ldots & \alpha_{p-1} & \cdot \\ l_1 & \ldots & l_{p-1} & l_p \end{bmatrix} = \bigcup_{\alpha_p} \begin{bmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{bmatrix}. \tag{5.2}$$

Each index α_p or l_p varies in some finite range. Since the explicit form of most of these ranges is irrelevant in the sequel, and since we do not want to overburden unnecessarily the notation, we shall not indicate these ranges explicitly. Rather, we just need

$$\begin{aligned} N\begin{bmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{bmatrix} &= \#\text{ roots in } \begin{bmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{bmatrix}, \\ N\begin{bmatrix} \alpha_1 & \ldots & \cdot \\ l_1 & \ldots & l_p \end{bmatrix} &= \#\text{ roots in } \begin{bmatrix} \alpha_1 & \ldots & \cdot \\ l_1 & \ldots & l_p \end{bmatrix}. \end{aligned} \tag{5.3}$$

Evidently, we have for $p>k$

$$\begin{aligned} \sum_{\alpha_p} N\begin{bmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{bmatrix} &= N\begin{bmatrix} \alpha_1 & \ldots & \alpha_{p-1} & \cdot \\ l_1 & \ldots & l_{p-1} & l_p \end{bmatrix}, \\ \sum_{\substack{\alpha_p,\ldots,\alpha_{k+1} \\ l_p,\ldots,l_{k+1}}} N\begin{bmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{bmatrix} &= N\begin{bmatrix} \alpha_1 & \ldots & \alpha_k \\ l_1 & \ldots & l_k \end{bmatrix}. \end{aligned} \tag{5.4}$$

(b) *The reduced Newton diagram of the phase.* Let $a_1<\ldots<a_l<\ldots<a_n$ be exponents so that the set of all roots of S'' different from $x\equiv0$, $y\equiv0$, can be divided as

$$\bigcup_{l=1}^{n} \begin{bmatrix} \cdot \\ l \end{bmatrix}.$$

Then we may write

$$S'' = U(x,y)x^{\alpha}y^{\beta}\prod_{l=1}^{n}\Phi\begin{bmatrix}\cdot\\ l\end{bmatrix}(x,y),$$
$$\Phi\begin{bmatrix}\cdot\\ l\end{bmatrix} \equiv \prod_{s}(y-r_s(x)), \quad 1\leqslant s\leqslant N\begin{bmatrix}\cdot\\ l\end{bmatrix}, \tag{5.5}$$

with $U(x,y)$ a factor bounded from above and below by positive constants. To alleviate the notation, we shall until the end of this section set

$$\alpha=\beta=0 \tag{5.6}$$

and indicate at the end how to adapt our arguments to the general case of arbitrary α,β. The following are the analogues of the quantities A_l, B_l introduced in the study of Model II:

$$A\begin{bmatrix}\cdot\\ l\end{bmatrix} = \sum_{\mu\leqslant l} a_\mu N\begin{bmatrix}\cdot\\ \mu\end{bmatrix}, \quad B\begin{bmatrix}\cdot\\ l\end{bmatrix} = \sum_{\mu\geqslant l+1} N\begin{bmatrix}\cdot\\ \mu\end{bmatrix}. \tag{5.7}$$

Observation 1. The vertices of the reduced Newton diagram are at the points

$$\left(1+A\begin{bmatrix}\cdot\\ l\end{bmatrix}, 1+B\begin{bmatrix}\cdot\\ l\end{bmatrix}\right), \quad 0\leqslant l\leqslant n. \tag{5.8}$$

The argument is parallel to the one given for the analogous observation in the study of Model II. We shall check that

(i) the points $\left(1+A\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right], 1+B\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]\right)$ do occur in the reduced Newton diagram;

(ii) the region above the sequence of line segments joining $\left(1+A\left[\begin{smallmatrix}\cdot\\ l-1\end{smallmatrix}\right], 1+B\left[\begin{smallmatrix}\cdot\\ l-1\end{smallmatrix}\right]\right)$ and $\left(1+A\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right], 1+B\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]\right)$ is convex;

(iii) the translate by $(1,1)$ of any other point (p,q) with x^py^q occurring with non-vanishing coefficient in the Taylor expansion of $S''(x,y)$ must lie in this region.

The first statement (i) is easy, since

$$y^{N\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]} \quad \text{and} \quad x^{a_l N\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]}$$

are respectively the highest power in y and the lowest power in x among the pure x-terms in the expansion of $\Phi\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]$. Carrying out the product in (5.5) for $S''(x,y)$ will produce then terms of the form

$$y^{B\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]}x^{A\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]}$$

with non-vanishing coefficients. Since the reduced Newton diagram of $S(x,y)$ is obtained by translating by $(1,1)$ the Newton diagram of $S''(x,y)$, the statement (i) follows.

The verification of (ii) is identical to the case of Model II. The equation of the segment joining $\left(1+A\left[{\cdot \atop l-1}\right], 1+B\left[{\cdot \atop l-1}\right]\right)$ and $\left(1+A\left[{\cdot \atop l}\right], 1+B\left[{\cdot \atop l}\right]\right)$ is just as before

$$q-\frac{1}{a_l}\left(-p+1+A\begin{bmatrix}\cdot\\ l\end{bmatrix}+\left(1+B\begin{bmatrix}\cdot\\ l\end{bmatrix}\right)a_l\right)=0. \tag{5.9}$$

Evaluating the right-hand side at $\left(1+A\left[{\cdot \atop l+1}\right], 1+B\left[{\cdot \atop l+1}\right]\right)$ gives

$$\left(\frac{a_{l+1}}{a_l}-1\right)N\begin{bmatrix}\cdot\\ l+1\end{bmatrix}>0,$$

which shows that $\left(1+A\left[{\cdot \atop l+1}\right], 1+B\left[{\cdot \atop l+1}\right]\right)$ lies strictly above the line joining $\left(1+A\left[{\cdot \atop l-1}\right], 1+B\left[{\cdot \atop l-1}\right]\right)$ and $\left(1+A\left[{\cdot \atop l}\right], 1+B\left[{\cdot \atop l}\right]\right)$, and thus that the region indicated is convex.

To verify (iii), we shall show how to locate all the points (p,q) with $x^p y^q$ occurring with non-zero coefficients in the expansion of $S''(x,y)$. Ignore first the higher-order terms in the expansion of $\Phi\left[{\cdot \atop l}\right]$ and write

$$\Phi\begin{bmatrix}\cdot\\ l\end{bmatrix}=y^{N\left[{\cdot \atop l}\right]}+\ldots+x^{a_l N\left[{\cdot \atop l}\right]}+\text{higher orders}, \tag{5.10}$$

where the dots stand for terms of the form

$$y^{N\left[{\cdot \atop l}\right]-k}x^{a_l\left(N\left[{\cdot \atop l}\right]+k\right)},\quad 1\leqslant k\leqslant N\begin{bmatrix}\cdot\\ l\end{bmatrix}-1$$

(some of which may have zero coefficients). We consider now the corresponding terms in the product (5.5) defining $S''(x,y)$, and isolate the term

$$y^{N\left[{\cdot \atop 2}\right]+\ldots+N\left[{\cdot \atop n}\right]}=y^{B\left[{\cdot \atop 1}\right]}$$

occurring in the product

$$\Phi\begin{bmatrix}\cdot\\ 2\end{bmatrix}\ldots\Phi\begin{bmatrix}\cdot\\ n\end{bmatrix}. \tag{5.11}$$

We note that this term, multiplied by the terms

$$y^{N\left[{\cdot \atop 1}\right]}\quad\text{and}\quad x^{a_1 N\left[{\cdot \atop 1}\right]}$$

of $\Phi\left[{\cdot \atop 1}\right]$, produces exactly the two end points $\left(0, B\left[{\cdot \atop 0}\right]\right)$ and $\left(a_1 N\left[{\cdot \atop 1}\right], B\left[{\cdot \atop 0}\right]-N\left[{\cdot \atop 1}\right]\right)=\left(A\left[{\cdot \atop 1}\right], B\left[{\cdot \atop 1}\right]\right)$ of the uppermost boundary segment of the Newton diagram of $S''(x,y)$. A key property of this boundary segment is that its slope is $-1/a_1$. We shall also say, informally, that it has "length" $N\left[{\cdot \atop 1}\right]$, since this is the amount by which the q-coordinate can change.

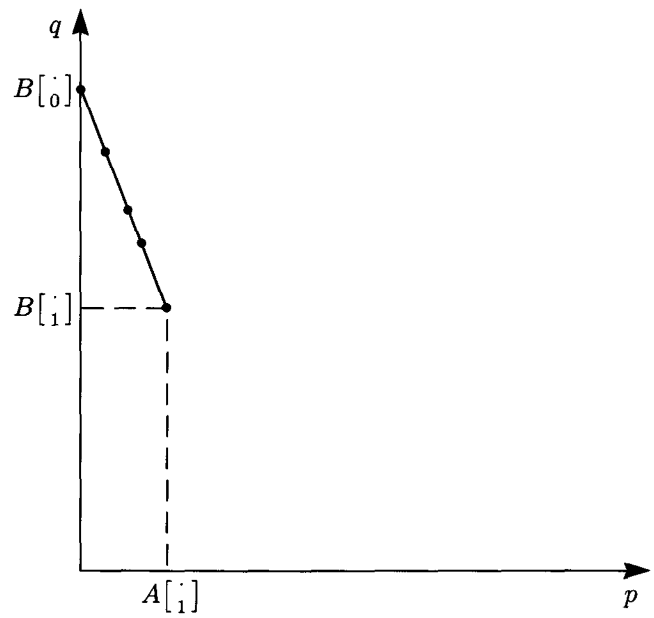

Fig. 1. The Newton diagram of S''. Vertices due to $\Phi\begin{bmatrix}\cdot\\1\end{bmatrix}y^{B\left[{\cdot \atop 1}\right]}$.

We can now identify easily the contributions of the intermediate terms given by the dots in (5.10), when multiplied by

$$y^{B\left[{\cdot \atop 1}\right]},$$

i.e., of the terms occurring in

$$\Phi\begin{bmatrix}\cdot\\1\end{bmatrix}y^{B\left[{\cdot \atop 1}\right]}=\Phi\begin{bmatrix}\cdot\\1\end{bmatrix}y^{N\left[{\cdot \atop 2}\right]+\dots+N\left[{\cdot \atop n}\right]}. \tag{5.12}$$

Each of these terms corresponds to an increase in the (p,q)-coordinates proportional to $(a_1,-1)$, and thus produces a point on the above boundary line segment. In particular, these points are in the convex hull of the vertices $\left(A\left[{\cdot \atop l}\right],B\left[{\cdot \atop l}\right]\right)$. There are at most $N\left[{\cdot \atop 1}\right]-1$ of them (Figure 1).

What are the effects of the intermediate terms in $\Phi\left[{\cdot \atop 2}\right]$ if we keep

$$y^{\left[{\cdot \atop 3}\right]+\dots+\left[{\cdot \atop n}\right]}$$

as the term arising from the product

$$\Phi\begin{bmatrix}\cdot\\3\end{bmatrix}\dots\Phi\begin{bmatrix}\cdot\\n\end{bmatrix}, \tag{5.13}$$

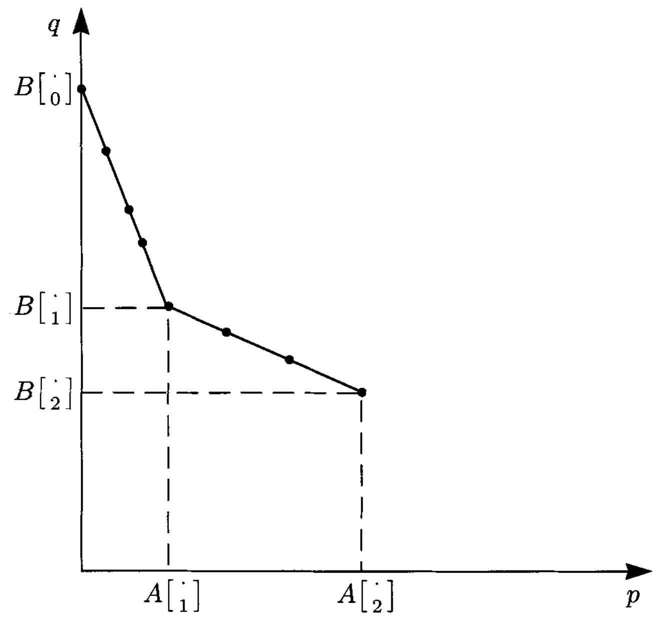

Fig. 2. The Newton diagram of S''. Vertices due to $x^{A\begin{bmatrix}\cdot\\1\end{bmatrix}}\Phi\begin{bmatrix}\cdot\\2\end{bmatrix}y^{B\begin{bmatrix}\cdot\\2\end{bmatrix}}$.

i.e., what are the points (p,q) arising from

$$\Phi\begin{bmatrix}\cdot\\1\end{bmatrix}\Phi\begin{bmatrix}\cdot\\2\end{bmatrix}y^{B\begin{bmatrix}\cdot\\2\end{bmatrix}}? \tag{5.14}$$

Consider first the terms

$$x^{a_1N\begin{bmatrix}\cdot\\1\end{bmatrix}}\Phi\begin{bmatrix}\cdot\\2\end{bmatrix}y^{B\begin{bmatrix}\cdot\\2\end{bmatrix}}$$

(Figure 2). We note that the extreme terms

$$y^{N\begin{bmatrix}\cdot\\2\end{bmatrix}} \quad \text{and} \quad x^{a_2N\begin{bmatrix}\cdot\\2\end{bmatrix}}$$

within $\Phi\begin{bmatrix}\cdot\\2\end{bmatrix}$ lead to the two end points of the next boundary line segment of the region defined by the convex hull of the points $\left(A\begin{bmatrix}\cdot\\l\end{bmatrix}, B\begin{bmatrix}\cdot\\l\end{bmatrix}\right)$. This boundary segment has slope $-1/a_2$, length $N\begin{bmatrix}\cdot\\2\end{bmatrix}$, and originates from the lower end point of the previous boundary segment. If we consider now the contributions of the intermediate terms in $\Phi\begin{bmatrix}\cdot\\2\end{bmatrix}$, we see that, just as in the previous case, they contribute points on the second boundary line segment.

Similarly, the terms in

$$y^{N\begin{bmatrix}\cdot\\1\end{bmatrix}}\Phi\begin{bmatrix}\cdot\\2\end{bmatrix}y^{B\begin{bmatrix}\cdot\\2\end{bmatrix}}$$

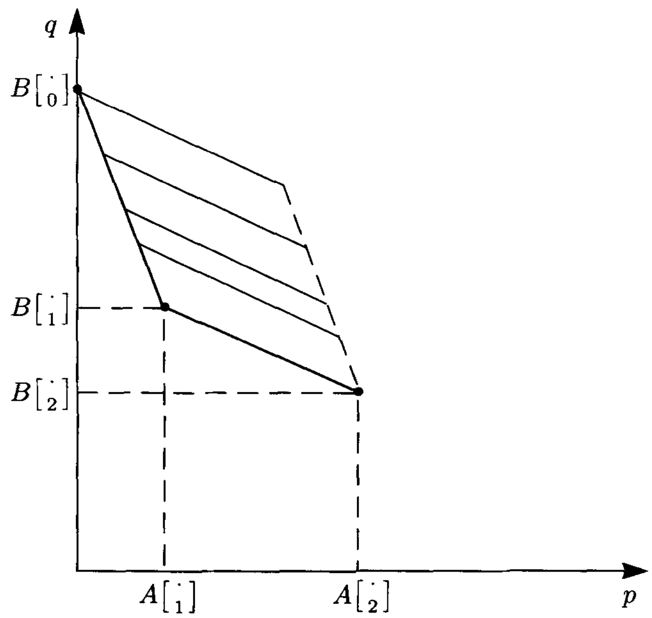

Fig. 3. The Newton diagram of S''. Vertices due to $\Phi\begin{bmatrix}\cdot\\1\end{bmatrix}\Phi\begin{bmatrix}\cdot\\2\end{bmatrix}y^{B\left[{\cdot \atop 2}\right]}$.

lead to points on the segment of slope $-1/a_2$, of length $N\left[{\cdot \atop 2}\right]$, and originating from the other end point of the first boundary segments, while the intermediate terms from

$$\Phi\begin{bmatrix}\cdot\\1\end{bmatrix}\Phi\begin{bmatrix}\cdot\\2\end{bmatrix}y^{B\left[{\cdot \atop 2}\right]}$$

can be located on parallel segments, of same length, but originating from the intermediate points in the first boundary segment. Clearly, they are all contained in the convex hull of the points $\left(A\left[{\cdot \atop l}\right], B\left[{\cdot \atop l}\right]\right)$ (Figure 3).

We can evidently continue in this manner, and locate all (m,n) arising from

$$\Phi\begin{bmatrix}\cdot\\1\end{bmatrix}\Phi\begin{bmatrix}\cdot\\2\end{bmatrix}\Phi\begin{bmatrix}\cdot\\3\end{bmatrix}y^{B\left[{\cdot \atop 3}\right]}$$

on the segments of length $N\left[{\cdot \atop 3}\right]$, slope $-1/a_3$, originating from any of the points obtained previously (Figure 4). Clearly, all these segments lie within the desired convex hull.

We have so far ignored higher-order terms in the expansion of $\Phi\left[{\cdot \atop l}\right]$. However, the above argument shows easily what is the effect of such terms: at each stage l, they are located on segments of length $N\left[{\cdot \atop l}\right]$, originating from the points obtained in previous stages, and of slopes greater than $-1/a_l$. Thus they lie well inside the convex hull of the points $\left(A\left[{\cdot \atop l}\right], B\left[{\cdot \atop l}\right]\right)$ (Figure 5).

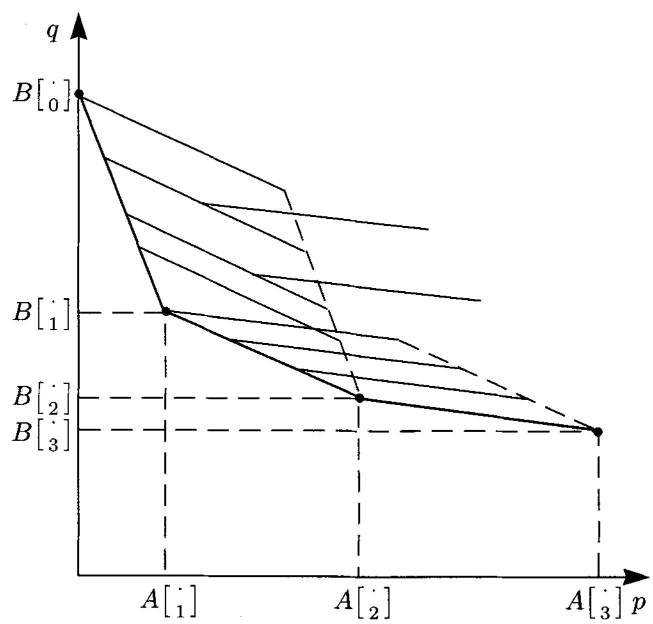

Fig. 4. The Newton diagram of S''. Vertices due to $\Phi\begin{bmatrix}\cdot\\1\end{bmatrix}\Phi\begin{bmatrix}\cdot\\2\end{bmatrix}\Phi\begin{bmatrix}\cdot\\3\end{bmatrix}y^{B\begin{bmatrix}\cdot\\3\end{bmatrix}}$.

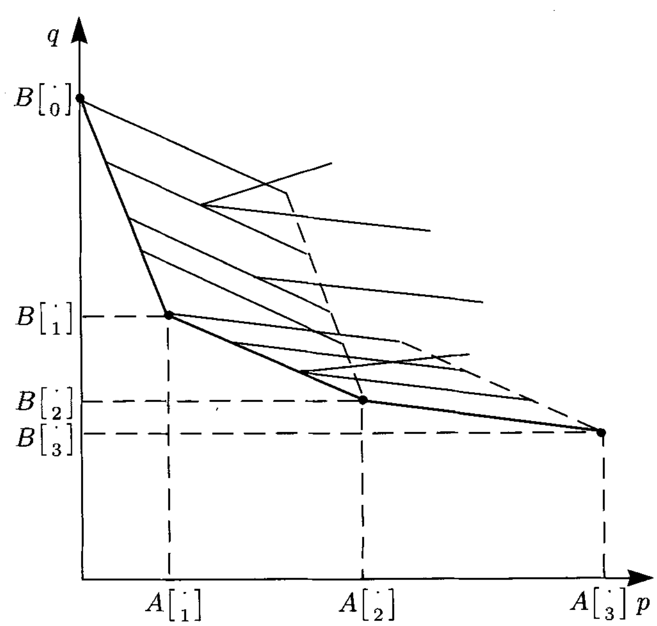

Fig. 5. The Newton diagram of S''.

Observation 2. It follows from our discussion that the leading exponents a_l of the roots of $S''(x,y)$ can be read off from its Newton diagram, together with their "generalized multiplicities" $N\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]$ (i.e., the number of roots with same leading exponent a_l).

Observation 3. The coordinates $(\delta_l^{-1},\delta_l^{-1})$ of the intersections of the boundary lines of the reduced Newton diagram of $S(x,y)$ with the bisectrix $p=q$ are given by

$$\delta_l=\frac{1+a_l}{1+A\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]+\left(1+B\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]\right)a_l} \tag{5.15}$$

as is readily seen from (5.9). We again set

$$\delta=\min_l \delta_l. \tag{5.16}$$

For future reference, we note that in presence of non-trivial factors $x^\alpha y^\beta$ in the factorization of $S''(x,y)$, the above discussion goes through unchanged, up to the shifts

$$A\begin{bmatrix}\cdot\\ l\end{bmatrix}\to A\begin{bmatrix}\cdot\\ l\end{bmatrix}+\alpha,\quad B\begin{bmatrix}\cdot\\ l\end{bmatrix}\to B\begin{bmatrix}\cdot\\ l\end{bmatrix}+\beta$$

in all the formulas. This completes our discussion of the reduced Newton diagram for $S(x,y)$.

(c) *Resolutions.* We now return to the proof of Theorem 1. Recall that we are for the moment restricting ourselves to resumming the operators T_{jk} in the range where

$$k>j-K,$$

where K is some large constant (cf. (3.5)–(3.8)), and that the Hessian $S''(x,y)$ has been factorized in (5.5). Continuing in this manner, we have, at each level of resolution p,

$$|S''|\sim U(x,y)\prod_{\substack{\alpha_1,\dots,\alpha_p\\ l_1,\dots,l_p}}\Phi\begin{bmatrix}\alpha_1 & \dots & \alpha_p\\ l_1 & \dots & l_p\end{bmatrix}(x,y). \tag{5.17}$$

Here as well as henceforth, we are using for the sake of simplicity the same notation for

$$\Phi\begin{bmatrix}\alpha_1 & \dots & \alpha_p\\ l_1 & \dots & l_p\end{bmatrix}(x,y)$$

and its absolute value. As we have seen in the treatment of complex roots (§4 (f)), we may assume without loss of generality that the roots $r(x)$ appearing in (5.5) are all Puiseux series with real coefficients. Clearly, since we are assuming that $k>j-K$, the ranges we consider below for k, namely $k\sim a_l j$ and $k\ll a_{l+1}j$, are only relevant when $a_l, a_{l+1}\geqslant 1$. Thus we may assume that $a_l, a_{l+1}\geqslant 1$ for the time being.

Consider now the usual dyadic decomposition $y\sim 2^{-k}$, $x\sim 2^{-j}$.

If $a_l j\ll k\ll a_{l+1}j$, then just as before

$$\Phi\begin{bmatrix} \cdot \\ \mu \end{bmatrix} \sim 2^{-ja_\mu N\left[\begin{smallmatrix} \cdot \\ \mu \end{smallmatrix}\right]}, \quad \mu\leqslant l,$$
$$\Phi\begin{bmatrix} \cdot \\ \mu \end{bmatrix} \sim 2^{-kN\left[\begin{smallmatrix} \cdot \\ \mu \end{smallmatrix}\right]}, \quad \mu\geqslant l+1,$$
$$|S''|\sim 2^{-j\left(a_1 N\left[\begin{smallmatrix} \cdot \\ 1 \end{smallmatrix}\right]+\ldots+a_l N\left[\begin{smallmatrix} \cdot \\ l \end{smallmatrix}\right]\right)} 2^{-k\left(N\left[\begin{smallmatrix} \cdot \\ l+1 \end{smallmatrix}\right]+\ldots+N\left[\begin{smallmatrix} \cdot \\ n \end{smallmatrix}\right]\right)}$$

and the arguments of Model II apply verbatim, e.g., $\|T\|\leqslant|\lambda|^{-\frac{1}{2}\Delta}$, with

$$\Delta\geqslant\frac{1+a_l}{1+a_1N\left[\begin{smallmatrix} \cdot \\ 1 \end{smallmatrix}\right]+\ldots+a_lN\left[\begin{smallmatrix} \cdot \\ l \end{smallmatrix}\right]+a_l\left(N\left[\begin{smallmatrix} \cdot \\ l+1 \end{smallmatrix}\right]+\ldots+N\left[\begin{smallmatrix} \cdot \\ n \end{smallmatrix}\right]+1\right)}\geqslant\delta.$$

Consider next the range $k\sim a_l j$ for some l. Then

$$\begin{aligned} &\Phi\begin{bmatrix} \cdot \\ \mu \end{bmatrix} \sim 2^{-ja_\mu N\left[\begin{smallmatrix} \cdot \\ \mu \end{smallmatrix}\right]}, \quad \mu<l, \\ &\Phi\begin{bmatrix} \cdot \\ \mu \end{bmatrix} \sim 2^{-ja_l}, \quad \mu>l, \end{aligned} \tag{5.18}$$

and we have

$$|S''|\sim 2^{-j\left(a_1N\left[\begin{smallmatrix} \cdot \\ 1 \end{smallmatrix}\right]+\ldots+a_{l-1}N\left[\begin{smallmatrix} \cdot \\ l-1 \end{smallmatrix}\right]\right)}\Phi\begin{bmatrix} \cdot \\ l \end{bmatrix} 2^{-ja_l\left(N\left[\begin{smallmatrix} \cdot \\ l+1 \end{smallmatrix}\right]+\ldots+N\left[\begin{smallmatrix} \cdot \\ n \end{smallmatrix}\right]\right)}. \tag{5.19}$$

Thus we are reduced to the case of a single exponent a_l, case similar to Model III, whose treatment we shall now follow.

The key observation is that, by going to finer and finer resolutions, superposing at each step a finer cut-off $\prod_{q=1}^{p}\chi_{m_q}(y-r_q(x))$, we will arrive after a finite number of steps at a resolution of S'' with the following properties:

(i) For each curved box in the final resolution, there is a resolution of a cluster of roots of S'',

$$\begin{bmatrix} \cdot \\ l_1 \end{bmatrix}\supset\begin{bmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{bmatrix}\supset\ldots\supset\begin{bmatrix} \alpha_1 & \alpha_2 & \ldots & \cdot \\ l_1 & l_2 & \ldots & l_p \end{bmatrix}\supset\begin{bmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_p \\ l_1 & l_2 & \ldots & l_p \end{bmatrix}, \tag{5.20}$$

such that the last resolution
$$\begin{bmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_p \\ l_1 & l_2 & \ldots & l_p \end{bmatrix}$$

contains only a single root (counted with its multiplicity);

(ii) On the curved box of the final resolution, the Hessian of S satisfies an estimate from above and below of the form

$$\begin{aligned}|S''| \sim\ & 2^{-j\mathcal{N}\left[\begin{smallmatrix} \cdot \\ l_1 \end{smallmatrix}\right]} 2^{-j\mathcal{N}\left[\begin{smallmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{smallmatrix}\right]} \dots 2^{-j\mathcal{N}\left[\begin{smallmatrix} \alpha_1 & \alpha_2 & \dots & \cdot \\ l_1 & l_2 & \dots & l_p \end{smallmatrix}\right]} \\ & \times 2^{ja_{l_1} N\left[\begin{smallmatrix} \alpha_1 \\ l_1 \end{smallmatrix}\right]} 2^{ja^{\alpha_1}_{l_1 l_2} N\left[\begin{smallmatrix} \alpha_1 & \alpha_2 \\ l_1 & l_2 \end{smallmatrix}\right]} \dots 2^{ja^{\alpha_1 \dots \alpha_{p-1}}_{l_1 \dots l_{p-1} l_p} N\left[\begin{smallmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_p \\ l_1 & l_2 & \dots & l_p \end{smallmatrix}\right]} \\ & \times \Phi\begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_p \\ l_1 & l_2 & \dots & l_p \end{bmatrix},\end{aligned} \tag{5.21}$$

where the positive numbers

$$\mathcal{N}\begin{bmatrix} \alpha_1 & \dots & \alpha_{p-1} & \cdot \\ l_1 & \dots & l_{p-1} & l_p \end{bmatrix}$$

are defined by

$$\begin{aligned}\mathcal{N}\begin{bmatrix} \alpha_1 & \dots & \alpha_{p-1} & \cdot \\ l_1 & \dots & l_{p-1} & l_p \end{bmatrix} = & \sum_{\mu \leqslant l_p} a^{\alpha_1 \dots \alpha_{p-1}}_{l_1 \dots l_{p-1} \mu} N\begin{bmatrix} \alpha_1 & \dots & a_{p-1} & \cdot \\ l_1 & \dots & l_{p-1} & \mu \end{bmatrix} \\ & + a^{\alpha_1 \dots \alpha_{p-1}}_{l_1 \dots l_{p-1} l_p} \sum_{\mu \geqslant l_p + 1} N\begin{bmatrix} \alpha_1 & \dots & a_{p-1} & \cdot \\ l_1 & \dots & l_{p-1} & \mu \end{bmatrix}.\end{aligned} \tag{5.22}$$

In the next subsection, we provide the algorithm leading to the resolutions satisfying (i) and (ii).

We also note the following special case of the quantities

$$\mathcal{N}\begin{bmatrix} \alpha_1 & \dots & \alpha_{p-1} & \cdot \\ l_1 & \dots & l_{p-1} & l_p \end{bmatrix}$$

introduced in (5.22)

$$\mathcal{N}\begin{bmatrix} \cdot \\ l \end{bmatrix} = A\begin{bmatrix} \cdot \\ l \end{bmatrix} + a_l B\begin{bmatrix} \cdot \\ l \end{bmatrix}. \tag{5.23}$$

(d) *Algorithm for the resolution of roots.* Returning to the decomposition T_{jk} in the range $k \sim a_l j$, we note that if the cluster $\left[\begin{smallmatrix} \cdot \\ l \end{smallmatrix}\right]$ consists of a single $\left[\begin{smallmatrix} \alpha \\ l \end{smallmatrix}\right]$, which itself contains only a single root counted with its multiplicity, then (5.19) shows that we have the desired resolution. This case becomes identical to that of Model II, where the root $y - x^{a_l}$ also does not resolve any further. The arguments there apply to give the desired estimate.

Thus we assume that

$$\Phi\begin{bmatrix} \cdot \\ l \end{bmatrix} = \prod_\gamma \Phi\begin{bmatrix} \gamma \\ l \end{bmatrix}.$$

The estimate (5.19) for the Hessian of the phase S can then be rewritten as

$$|S''|\sim 2^{-jN\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]}\prod_{\gamma}2^{ja_lN\left[\begin{smallmatrix}\gamma\\ l\end{smallmatrix}\right]}\Phi\begin{bmatrix}\gamma\\ l\end{bmatrix}. \tag{5.24}$$

We note that the above estimate (5.24) is of the desired form (5.21), only with possibly many factors Φ on the right-hand side, instead of a single

$$\Phi\begin{bmatrix}\alpha_1 & \dots & \alpha_p\\ l_1 & \dots & l_p\end{bmatrix}$$

as in (5.21). Thus our first main task is to decrease the number of γ's on the right-hand side of (5.24).

We select an index α among the γ's. Let $r\left[\begin{smallmatrix}\alpha\\ l\end{smallmatrix}\right](x)$ be a root within $\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]$ with highest exponent a^{α}_{lm}, and set

$$y-r\begin{bmatrix}\alpha\\ l\end{bmatrix}(x)\sim 2^{-m},\quad m\geqslant a_lj.$$

We need to consider again several different cases.

If $m\gg a_lj$, then all the factors $\Phi\left[\begin{smallmatrix}\gamma\\ l\end{smallmatrix}\right]$ are for $\gamma\neq\alpha$ of size

$$\Phi\begin{bmatrix}\gamma\\ l\end{bmatrix}\sim 2^{-a_ljN\left[\begin{smallmatrix}\gamma\\ l\end{smallmatrix}\right]},\quad \gamma\neq\alpha.$$

It follows at once that

$$|S''|\sim 2^{-jN\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]}2^{ja_lN\left[\begin{smallmatrix}\alpha\\ l\end{smallmatrix}\right]}\Phi\begin{bmatrix}\alpha\\ l\end{bmatrix}. \tag{5.25}$$

This is again of the form (5.21), but with a single factor $\Phi\left[\begin{smallmatrix}\alpha\\ l\end{smallmatrix}\right]$.

If $m\sim a_lj$, then the estimate (5.24) for S'' reduces at once to

$$|S''|\sim 2^{-jN\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]}\prod_{\gamma\neq\alpha}2^{ja_{l_1}N\left[\begin{smallmatrix}\gamma\\ l\end{smallmatrix}\right]}\Phi\begin{bmatrix}\gamma\\ l\end{bmatrix}, \tag{5.26}$$

which is of the form (5.24), but with one less factor among the $\Phi\left[\begin{smallmatrix}\gamma\\ l\end{smallmatrix}\right]$'s.

Thus we decrease the number of factors Φ in all eventualities. In the case $m\gg a_lj$, we already arrived at a single $\Phi\left[\begin{smallmatrix}\alpha\\ l\end{smallmatrix}\right]$. In the case $m\sim a_lj$, and if there are more than a single γ left which are different from α, we can repeat the process until there is also a single Φ left on the right-hand side of (5.26). In this way, by a single resolution, we have reached a stage where $|S''|$ satisfies estimates of the form (5.21), with however the condition (i), namely that the lone cluster Φ appearing on the right-hand side consists of a single root with its multiplicities, still possibly not satisfied.

Our next step is to introduce further resolutions, when necessary, to fulfill the condition (i) as well. This is achieved simply by repeating the resolution process. Now we started with a resolution $\prod_l \Phi\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right]$, into clusters of roots distinguished by their leading exponents

$$a_1 < \ldots < a_l < a_{l+1} < \ldots .$$

The stage described by (5.25) is the one where we have isolated a cluster of roots $c^\alpha x^{a_l} + \ldots$ among all the roots $c^\gamma x^{a_l} + \ldots$ with leading exponent x^{a_l}. Within the cluster $\left[\begin{smallmatrix}\alpha\\ l\end{smallmatrix}\right]$, there are smaller clusters of roots, distinguished by their next leading exponents after $c^\alpha x^{a_l}$. Indeed, set $l \equiv l_1$, $\alpha \equiv \alpha_1$. Then

$$\begin{bmatrix}\alpha_1\\ l_1\end{bmatrix} = \bigcup_{l_2} \begin{bmatrix}\alpha_1 & \cdot\\ l_1 & l_2\end{bmatrix} \tag{5.27}$$

and

$$\bigcup_{l_2} \begin{bmatrix}\alpha_1 & \cdot\\ l_1 & l_2\end{bmatrix} \ni r(x) \quad \Leftrightarrow \quad r(x) = c^{\alpha_1} x^{a_{l_1}} + c^{\alpha_1\alpha_2}_{l_1 l_2} x^{a^{\alpha_1}_{l_1 l_2}} + \ldots . \tag{5.28}$$

One cycle of our resolution process will be complete when we shall have reduced estimates for S'', originally written in terms of $\prod_l \Phi\left[\begin{smallmatrix}\cdot\\ l\end{smallmatrix}\right] \equiv \left[\begin{smallmatrix}\cdot\\ l_1\end{smallmatrix}\right]$, into estimates for S'', written in terms of

$$\begin{bmatrix}\alpha_1 & \cdot\\ l_1 & l_2\end{bmatrix},$$

for some fixed l_2.

More precisely, consider the range $k \sim a_l j$ of the decomposition of T into T_{jk}, and assume that we have already reduced S'' to an estimate of the form (5.25), with α, l denoted now by α_1, l_1, as in (5.27).

Then if $\left[\begin{smallmatrix}\alpha_1\\ l_1\end{smallmatrix}\right]$ consists of only one root counted with its multiplicity, we are done. Otherwise, consider the finer resolution of (5.27), and order the exponents $a^{\alpha_1}_{l_1 l_2}$ in increasing order in l_2,

$$a^{\alpha_1}_{l_1 \mu} < a^{\alpha_1}_{l_1(\mu+1)}. \tag{5.29}$$

Select a root $r(x)$ in

$$\bigcup_{l_2} \begin{bmatrix}\alpha_1 & \cdot\\ l_1 & l_2\end{bmatrix}$$

with highest exponent $a^{\alpha_1}_{l_1 l_2}$, and set

$$y - r(x) \sim 2^{-m},$$

where, to lighten the notation, we still use the index m, although it is distinct from the index m of earlier resolutions.

If $a_{l_1 l_2}^{\alpha_1} j \ll m \ll a_{l_1(l_2+1)}^{\alpha_1} j$, then

$$\Phi\begin{bmatrix}\alpha_1 & \cdot \\ l_1 & \mu\end{bmatrix} \sim 2^{-j a_{l_1\mu}^{\alpha_1} N\left[\begin{smallmatrix}\alpha_1 & \cdot \\ l_1 & \mu\end{smallmatrix}\right]}, \quad \mu \leqslant l_2,$$

$$\Phi\begin{bmatrix}\alpha_1 & \cdot \\ l_1 & \mu\end{bmatrix} \sim 2^{-m N\left[\begin{smallmatrix}\alpha_1 & \cdot \\ l_1 & \mu\end{smallmatrix}\right]}, \quad \mu \geqslant l_2+1.$$

It follows that $|S''|$ can be bounded by

$$|S''| \sim 2^{-j\mathcal{N}\left[\begin{smallmatrix}\cdot \\ l_1\end{smallmatrix}\right]} 2^{-j\left(\sum_{\mu \leqslant l_2} a_{l_1\mu}^{\alpha_1} N\left[\begin{smallmatrix}\alpha_1 & \cdot \\ l_1 & \mu\end{smallmatrix}\right] - a_{l_1} N\left[\begin{smallmatrix}\alpha_1 \\ l_1\end{smallmatrix}\right]\right)} 2^{-m\sum_{\mu \geqslant l_1+1} N\left[\begin{smallmatrix}\alpha_1 & \cdot \\ l_1 & \mu\end{smallmatrix}\right]}. \tag{5.30}$$

The size estimate of the corresponding operator is still $(\Delta x \Delta y)^{1/2} \leqslant 2^{-m} 2^{-\frac{1}{2}(1-a_l)j}$. Combining this with the oscillating estimate resulting from (5.30), we find the following decay rate $|\lambda|^{-\frac{1}{2}\Delta}$, after applying the summation Lemma 3,

$$\Delta = \frac{1+2a_{l_1 l_2}^{\alpha_1} - a_{l_1}}{1+\mathcal{N}\begin{bmatrix}\cdot \\ l_1\end{bmatrix} - a_{l_1} N\begin{bmatrix}\alpha_1 \\ l_1\end{bmatrix} + \mathcal{N}\begin{bmatrix}\alpha_1 & \cdot \\ l_1 & l_2\end{bmatrix} + 2a_{l_1 l_2}^{\alpha_1} - a_{l_1}}. \tag{5.31}$$

If we view the right-hand side of (5.31) as a function of $a_{l_1 l_2}^{\alpha_1}$, it is an increasing function (cf. Lemma 4), and thus its value is greater than the value we get by letting $a_{l_1 l_2}^{\alpha_1} \to a_{l_1(l_2-1)}^{\alpha_1}$. Continuing in this manner, we obtain

$$\begin{aligned}\Delta &\geqslant \frac{1+2a_{l_1(l_2-1)}^{\alpha_1} - a_{l_1}}{1+\mathcal{N}\begin{bmatrix}\cdot \\ l_1\end{bmatrix} - a_{l_1} N\begin{bmatrix}\alpha_1 \\ l_1\end{bmatrix} + \mathcal{N}\begin{bmatrix}\alpha_1 & \cdot \\ l_1 & l_2-1\end{bmatrix} + 2a_{l_1(l_2-1)}^{\alpha_1} - a_{l_1}} \\ &\geqslant \frac{1+2a_{l_1 1}^{\alpha_1} - a_{l_1}}{1+\mathcal{N}\begin{bmatrix}\cdot \\ l_1\end{bmatrix} - a_{l_1} N\begin{bmatrix}\alpha_1 \\ l_1\end{bmatrix} + \left(N\begin{bmatrix}\alpha_1 \\ l_1\end{bmatrix} + 2\right) a_{l_1 1}^{\alpha_1} - a_{l_1}}.\end{aligned} \tag{5.32}$$

Again using the fact that the right-hand side of (5.32) is an increasing function of $a_{l_1 1}^{\alpha_1}$, we can let $a_{l_1 1}^{\alpha_1}$ tend to a_{l_1}, and arrive at

$$\Delta \geqslant \frac{1+a_{l_1}}{1+\mathcal{N}\begin{bmatrix}\cdot \\ l_1\end{bmatrix} + a_{l_1}} = \delta_{l_1} \geqslant \delta$$

in view of (5.15), (5.16) and (5.23).

This shows that we need study only the operators T_{jk}^m arising from $m \sim a_{l_1 l_2}^{\alpha_1} j$. The number of such m's is boundedly finite, and it suffices to establish the desired rate for each m. In this case,

$$\begin{aligned}\Phi\begin{bmatrix}\alpha_1 & \cdot \\ l_1 & \mu\end{bmatrix} &\sim 2^{-j a_{l_1\mu}^{\alpha_1} N\left[\begin{smallmatrix}\alpha_1 & \cdot \\ l_1 & \mu\end{smallmatrix}\right]}, \quad \mu \leqslant l_1-1, \\ \Phi\begin{bmatrix}\alpha_1 & \cdot \\ l_1 & \mu\end{bmatrix} &\sim 2^{-j a_{l_1 l_2}^{\alpha_1} N\left[\begin{smallmatrix}\alpha_1 & \cdot \\ l_1 & \mu\end{smallmatrix}\right]}, \quad \mu \geqslant l_1+1.\end{aligned} \tag{5.33}$$

Thus the Hessian of the phase satisfies the following estimate from above and below

$$|S''| \sim 2^{-j\mathcal{N}\left[\begin{smallmatrix} \cdot \\ l_1 \end{smallmatrix}\right]} 2^{-j\mathcal{N}\left[\begin{smallmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{smallmatrix}\right]} 2^{ja_{l_1} N\left[\begin{smallmatrix} \alpha_1 \\ l_1 \end{smallmatrix}\right]} 2^{ja^{\alpha_1}_{l_1 l_2} N\left[\begin{smallmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{smallmatrix}\right]} \Phi\begin{bmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{bmatrix}. \tag{5.34}$$

This is an estimate of the form that we wanted, i.e., in terms of clusters

$$\Phi\begin{bmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{bmatrix}$$

distinguished by their next leading exponents $a^{\alpha_1}_{l_1 l_2}$.

A cycle of our resolution process is now complete. Evidently each cycle decreases the number of roots involved by at least one, so that, by repeating the resolution process a finite number of times, we can reach the stage where the condition (i) is also satisfied.

(e) *Estimates for the operators of the resolution.* We shall now show that bounds of the form (5.21) lead to the desired decay rate for T. From the resolution process, at each step of further decompositions T^m_{jk} of T_{jk} given by $|y-r(x)| \sim 2^{-m}$ for some root $r(x)$, we observe that T^m_{jk} can be summed back to give the desired estimate for T_{jk} except for a boundedly finite number of values of m, clustered around some exponent $a^{\alpha_1 \ldots \alpha_{p-1}}_{l_1 \ldots l_p} j$. This statement is the exact analogue of the fact that, in the original decomposition of T into T_{jk}, the ranges of k not clustered around $a_l j$ for some exponent l can immediately be summed to give the desired result. Thus we need consider only each individual m as we go to the next resolution, and establish uniform individual bounds for the corresponding operator T^M_{jk}, when we arrive at the final resolution satisfying the conditions (i) and (ii).

Assume then (i) and (ii), and set

$$y - r(x) \sim 2^{-m}, \quad m \geqslant a^{\alpha_1 \ldots \alpha_{p-1}}_{l_1 \ldots l_p},$$

with $r(x)$ the only root left in

$$\begin{bmatrix} \alpha_1 & \ldots & \alpha_{p-1} & \cdot \\ l_1 & \ldots & l_{p-1} & l_p \end{bmatrix}.$$

Denote the resulting operator decomposition simply by T^m. Then the size and oscillatory estimates are given by

$$\begin{aligned} &\|T^m\| \leqslant 2^{-m} 2^{\frac{1}{2}(1-a_{l_1})j}, \\ &\|T^m\| \leqslant |\lambda|^{\frac{1}{2}} 2^{\frac{1}{2}j\left(\mathcal{N}\left[\begin{smallmatrix} \cdot \\ l_1 \end{smallmatrix}\right] + \mathcal{N}\left[\begin{smallmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{smallmatrix}\right] + \ldots + \mathcal{N}\left[\begin{smallmatrix} \alpha_1 & \ldots & \alpha_{p-1} & \cdot \\ l_1 & \ldots & l_{p-1} & l_p \end{smallmatrix}\right]\right)} \\ &\qquad \times 2^{-\frac{1}{2}j\left(a_{l_1} N\left[\begin{smallmatrix} \alpha_1 \\ l_1 \end{smallmatrix}\right] + a^{\alpha_1}_{l_1 l_2} N\left[\begin{smallmatrix} \alpha_1 & \alpha_2 \\ l_1 & l_2 \end{smallmatrix}\right] + \ldots + a^{\alpha_1 \ldots \alpha_{p-1}}_{l_1 \ldots l_p} N\left[\begin{smallmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{smallmatrix}\right]\right)} \\ &\qquad \times 2^{\frac{1}{2}mN\left[\begin{smallmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{smallmatrix}\right]}. \end{aligned} \tag{5.35}$$

In view of Lemma 3, we have

$$\sum_{m \geqslant a_{l_1 \ldots l_p}^{\alpha_1 \ldots \alpha_{p-1}}} \|T^m\| \leqslant |\lambda|^{-\frac{1}{2}\Delta}$$

with Δ=Numerator/Denominator,

$$\text{Numerator} = 1+2a_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_{p-1}} - a_{l_1},$$

$$\begin{aligned}\text{Denominator} = {} & 1+\mathcal{N}\begin{bmatrix} \cdot \\ l_1 \end{bmatrix} + \mathcal{N}\begin{bmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{bmatrix} + \ldots + \mathcal{N}\begin{bmatrix} \alpha_1 & \ldots & \alpha_{p-1} & \cdot \\ l_1 & \ldots & l_{p-1} & l_p \end{bmatrix} \\ & -\left(a_{l_1} N\begin{bmatrix} \alpha_1 \\ l_1 \end{bmatrix} + a_{l_1 l_2}^{\alpha_1} N\begin{bmatrix} \alpha_1 & \alpha_2 \\ l_1 & l_2 \end{bmatrix} + \ldots + a_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_{p-1}} N\begin{bmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{bmatrix}\right) \\ & + \left(N\begin{bmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{bmatrix} + 2\right) a_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_{p-1}} - a_{l_1}. \end{aligned} \tag{5.36}$$

We shall show that this rate is better than the rate δ. In view of Lemma 4, the fraction on the right-hand side of (5.36) is an increasing function of $a_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_{p-1}}$. Thus we get a lower bound by letting $a_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_{p-1}}$ decrease to $a_{l_1\ldots(l_p-1)}^{\alpha_1\ldots\alpha_{p-1}}$. We repeat the decreasing process until we reach $a_{l_1\ldots l_{p-1}1}^{\alpha_1\ldots\alpha_{p-1}}$. Evidently, the numerator of (5.36) becomes

$$1+2a_{l_1\ldots l_{p-1}1}^{\alpha_1\ldots\alpha_{p-1}} - a_{l_1}.$$

As for the denominator, since

$$\mathcal{N}\begin{bmatrix} \alpha_1 & \ldots & \alpha_{p-1} & \cdot \\ l_1 & \ldots & l_{p-1} & l_p \end{bmatrix} \downarrow \mathcal{N}\begin{bmatrix} \alpha_1 & \ldots & \alpha_{p-1} & \cdot \\ l_1 & \ldots & l_{p-1} & l_p - 1 \end{bmatrix}$$

as $a_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_{p-1}} \downarrow a_{l_1\ldots(l_p-1)}^{\alpha_1\ldots\alpha_{p-1}}$, it evolves successively as

$$\begin{aligned} & 1+\mathcal{N}\begin{bmatrix} \cdot \\ l_1 \end{bmatrix} + \mathcal{N}\begin{bmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{bmatrix} + \ldots + \mathcal{N}\begin{bmatrix} \alpha_1 & \ldots & \alpha_{p-1} & \cdot \\ l_1 & \ldots & l_{p-1} & l_p \end{bmatrix} \\ & -\left(a_{l_1} N\begin{bmatrix} \alpha_1 \\ l_1 \end{bmatrix} + a_{l_1 l_2}^{\alpha_1} N\begin{bmatrix} \alpha_1 & \alpha_2 \\ l_1 & l_2 \end{bmatrix} + \ldots + a_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_{p-1}} N\begin{bmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{bmatrix}\right) \\ & + \left(N\begin{bmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{bmatrix} + 2\right) a_{l_1\ldots l_p}^{\alpha_1\ldots\alpha_{p-1}} - a_{l_1} \\ & \qquad \to 1+\mathcal{N}\begin{bmatrix} \cdot \\ l_1 \end{bmatrix} + \mathcal{N}\begin{bmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{bmatrix} + \ldots + \mathcal{N}\begin{bmatrix} \alpha_1 & \ldots & \alpha_{p-1} & \cdot \\ l_1 & \ldots & l_{p-1} & l_p - 1 \end{bmatrix} \\ & \qquad -\left(a_{l_1} N\begin{bmatrix} \alpha_1 \\ l_1 \end{bmatrix} + a_{l_1 l_2}^{\alpha_1} N\begin{bmatrix} \alpha_1 & \alpha_2 \\ l_1 & l_2 \end{bmatrix} + \ldots + a_{l_1\ldots(l_p-1)}^{\alpha_1\ldots\alpha_{p-1}} N\begin{bmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{bmatrix}\right) \\ & \qquad + \left(N\begin{bmatrix} \alpha_1 & \ldots & \alpha_p \\ l_1 & \ldots & l_p \end{bmatrix} + 2\right) a_{l_1\ldots(l_p-1)}^{\alpha_1\ldots\alpha_{p-1}} - a_{l_1} \end{aligned}$$

and ultimately, when we have decreased $a_{l_1\dots l_p}^{\alpha_1\dots\alpha_{p-1}}$ all the way down to $a_{l_1\dots l_{p-1}1}^{\alpha_1\dots\alpha_{p-1}}$, as

$$\begin{aligned}
&1+\mathcal{N}\begin{bmatrix} \cdot \\ l_1 \end{bmatrix}+\mathcal{N}\begin{bmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{bmatrix}+\dots+\mathcal{N}\begin{bmatrix} \alpha_1 & \dots & \alpha_{p-2} & \cdot \\ l_1 & \dots & l_{p-2} & l_{p-1} \end{bmatrix}\\
&+a_{l_1\dots l_{p-1}1}^{\alpha_1\dots\alpha_{p-1}}N\begin{bmatrix} \alpha_1 & \dots & \alpha_{p-1} \\ l_1 & \dots & l_{p-1} \end{bmatrix}\\
&-\Bigg(a_{l_1}N\begin{bmatrix} \alpha_1 \\ l_1 \end{bmatrix}+a_{l_1l_2}^{\alpha_1}N\begin{bmatrix} \alpha_1 & \alpha_2 \\ l_1 & l_2 \end{bmatrix}+\dots+a_{l_1\dots l_{p-1}}^{\alpha_1\dots\alpha_{p-2}}N\begin{bmatrix} \alpha_1 & \dots & \alpha_{p-1} \\ l_1 & \dots & l_{p-1} \end{bmatrix}\\
&\qquad+a_{l_1\dots l_{p-1}1}^{\alpha_1\dots\alpha_{p-1}}N\begin{bmatrix} \alpha_1 & \dots & \alpha_p \\ l_1 & \dots & l_p \end{bmatrix}\Bigg)\\
&+\left(N\begin{bmatrix} \alpha_1 & \dots & \alpha_p \\ l_1 & \dots & l_p \end{bmatrix}+2\right)a_{l_1\dots l_{p-1}1}^{\alpha_1\dots\alpha_{p-1}}-a_{l_1}.
\end{aligned}\tag{5.37}$$

This last expression can be rewritten as

$$\begin{aligned}
&1+\mathcal{N}\begin{bmatrix} \cdot \\ l_1 \end{bmatrix}+\mathcal{N}\begin{bmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{bmatrix}+\dots+\mathcal{N}\begin{bmatrix} \alpha_1 & \dots & \alpha_{p-2} & \cdot \\ l_1 & \dots & l_{p-2} & l_{p-1} \end{bmatrix}\\
&-\left(a_{l_1}N\begin{bmatrix} \alpha_1 \\ l_1 \end{bmatrix}+a_{l_1l_2}^{\alpha_1}N\begin{bmatrix} \alpha_1 & \alpha_2 \\ l_1 & l_2 \end{bmatrix}+\dots+a_{l_1\dots l_{p-1}}^{\alpha_1\dots\alpha_{p-2}}N\begin{bmatrix} \alpha_1 & \dots & \alpha_{p-1} \\ l_1 & \dots & l_{p-1} \end{bmatrix}\right)\\
&+\left(N\begin{bmatrix} \alpha_1 & \dots & \alpha_{p-1} \\ l_1 & \dots & l_{p-1} \end{bmatrix}+2\right)a_{l_1\dots l_{p-1}1}^{\alpha_1\dots\alpha_{p-1}}-a_{l_1}.
\end{aligned}$$

If we decrease now $a_{l_1\dots l_{p-1}1}^{\alpha_1\dots\alpha_{p-1}}$ to $a_{l_1\dots l_{p-2}l_{p-1}}^{\alpha_1\dots\alpha_{p-2}}$, the preceding expression becomes

$$\begin{aligned}
&1+\mathcal{N}\begin{bmatrix} \cdot \\ l_1 \end{bmatrix}+\mathcal{N}\begin{bmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{bmatrix}+\dots+\mathcal{N}\begin{bmatrix} \alpha_1 & \dots & \alpha_{p-2} & \cdot \\ l_1 & \dots & l_{p-2} & l_{p-1} \end{bmatrix}\\
&-\left(a_{l_1}N\begin{bmatrix} \alpha_1 \\ l_1 \end{bmatrix}+a_{l_1l_2}^{\alpha_1}N\begin{bmatrix} \alpha_1 & \alpha_2 \\ l_1 & l_2 \end{bmatrix}+\dots+a_{l_1\dots l_{p-1}}^{\alpha_1\dots\alpha_{p-2}}N\begin{bmatrix} \alpha_1 & \dots & \alpha_{p-1} \\ l_1 & \dots & l_{p-1} \end{bmatrix}\right)\\
&+\left(N\begin{bmatrix} \alpha_1 & \dots & \alpha_{p-1} \\ l_1 & \dots & l_{p-1} \end{bmatrix}+2\right)a_{l_1\dots l_{p-2}l_{p-1}}^{\alpha_1\dots\alpha_{p-2}}-a_{l_1},
\end{aligned}$$

which is exactly the same denominator we started with, but with p replaced by $p-1$. We can thus decrease p all the way to $p=2$, and find for the denominator

$$\begin{aligned}
&1+\mathcal{N}\begin{bmatrix} \cdot \\ l_1 \end{bmatrix}+\mathcal{N}\begin{bmatrix} \alpha_1 & \cdot \\ l_1 & l_2 \end{bmatrix}-\left(a_{l_1}N\begin{bmatrix} \alpha_1 \\ l_1 \end{bmatrix}+a_{l_1l_2}^{\alpha_1}N\begin{bmatrix} \alpha_1 & \alpha_2 \\ l_1 & l_2 \end{bmatrix}\right)\\
&+\left(N\begin{bmatrix} \alpha_1 & \alpha_2 \\ l_1 & l_2 \end{bmatrix}+2\right)a_{l_1l_2}^{\alpha_1}-a_{l_1},
\end{aligned}$$

which reduces to, upon decreasing $a_{l_1l_2}^{\alpha_1}$ to a_{l_1},

$$1+\mathcal{N}\begin{bmatrix} \cdot \\ l_1 \end{bmatrix}-a_{l_1}N\begin{bmatrix} \alpha_1 \\ l_1 \end{bmatrix}+\left(N\begin{bmatrix} \alpha_1 \\ l_1 \end{bmatrix}+2\right)a_{l_1}-a_{l_1}=1+\mathcal{N}\begin{bmatrix} \cdot \\ l_1 \end{bmatrix}+a_{l_1}.\tag{5.38}$$

Since the denominator has become $1+a_{l_1}$ in the process, we can conclude that

$$\delta \geqslant \frac{1+a_{l_1}}{1+\mathcal{N}\begin{bmatrix} \cdot \\ l_1 \end{bmatrix}+a_{l_1}} = \delta_l. \tag{5.39}$$

The bounds for T_{jk} in the range $k>j-K$ follow.

We provide now the modifications we need in presence of factors of the form $x^\alpha y^\beta$ in the factorization of $S''(s,y)$. In the range $a_l j \ll k \ll a_{l+1} j$, simply shift in all our arguments $A\left[\begin{smallmatrix} \cdot \\ l \end{smallmatrix}\right]$ by α, $B\left[\begin{smallmatrix} \cdot \\ l \end{smallmatrix}\right]$ by β. In particular, the denominator in the expression (5.15) for δ_l shifts by $\alpha+\beta a_l$, which gives the decay rate we want. The previous arguments apply verbatim. In the range $k \sim a_l j$, we note that the size of $S''(x,y)$ is then modified by a constant prefactor of $2^{-(\alpha+\beta a_l)j}$, everything else remaining the same. As we can see from Lemma 3, the net effect is again to shift the final estimate in terms of δ_l by the additional $(\alpha+\beta a_l)$-term in the denominator just as before.

To complete the proof of Theorem 1, we consider the remaining range $k \leqslant j-K$ in the decomposition (3.5). However, in this range, it suffices to write the Hessian $S''(x,y)$ as a polynomial in x, with analytic coefficients in y (up to the usual non-vanishing factor). The zeroes of S'' are then of the form $x=\tilde{r}_s(y)$, with $\tilde{r}_s(y)$ Puiseux series with leading exponents b_s. We can now repeat our arguments, with the roles of x and y interchanged. Evidently, the range $j \geqslant k+K$ we are now considering can only fit in $j \sim b_s k$, or $j \ll b_s k$ for $b_s>1$. Hence the above argument applies, and the proof of (1.3) is complete.

To prove the converse inequality, we consider three cases depending on whether the main face of the reduced Newton diagram when extended intersects both the p- and q-axes; or whether the face is parallel to one of the axes; or lastly whether the face reduces to a single vertex. In the first case, let $(\alpha,0)$ and $(0,\beta)$ denote the intersections with the p- and q-axes respectively. Then

$$\delta = \frac{1}{\alpha}+\frac{1}{\beta}.$$

Now for large positive λ, define the functions f_λ, g_λ by

$$f_\lambda(y) = \begin{cases} 1 & \text{if } 1 \leqslant \lambda y^\beta \leqslant 1+c_1, \\ 0 & \text{otherwise,} \end{cases}$$

$$g_\lambda(x) = \begin{cases} 1 & \text{if } 1 \leqslant \lambda x^\alpha \leqslant 1+c_1, \\ 0 & \text{otherwise.} \end{cases}$$

Here c_1 is a small constant to be fixed later. Note that $\|f_\lambda\| \sim \lambda^{-\frac{1}{2\beta}}$ and $\|g_\lambda\| \sim \lambda^{-\frac{1}{2\alpha}}$ as $\lambda \to \infty$.

With $S(x,y)=\sum_{p,q} c_{pq}x^p y^q$, let $S_0=\sum' c_{pq}$, when the summation is restricted to the main face of the reduced Newton diagram. Then, as is easily verified, for any $\varepsilon>0$, in the support of $f_\lambda(y)g_\lambda(x)$, we have

$$|\lambda S(x,y)-S_0|<\varepsilon,$$

as long as c_1 is taken to be small in terms of $\sum' |c_{pq}|$, and then λ is taken to be large. If we take, say $\varepsilon<\frac{1}{2}\pi$, then this shows that

$$|(T_\lambda f_\lambda, g_\lambda)| \geqslant c\left(\int g_\lambda(x)\,dx\right)\left(\int f_\lambda(y)\,dy\right) \sim \lambda^{-\frac{1}{\alpha}}\lambda^{-\frac{1}{\beta}}$$

as $\lambda\to\infty$. Hence

$$\frac{|(T_\lambda f_\lambda, g_\lambda)|}{\|f_\lambda\|\cdot\|g_\lambda\|} \sim \lambda^{-\frac{1}{2}\left(\frac{1}{\alpha}+\frac{1}{\beta}\right)},$$

proving our claim. The case when the main face is parallel to the p-axis corresponds effectively to the above situation when $\alpha=\infty$. In this situation, we define g_λ to be independent of λ and to be the characteristic function of a small interval around the origin. The argument then proceeds as before; the situation is also similar when the main face is parallel to the q-axis. Finally, when the main face reduces to a vertex (c,c), we take $\alpha=\beta=2c$ and argue as before.

6. Further remarks

We discuss briefly some closely related developments.

(a) We have recently shown that the methods of this paper can also be adapted to generalize Theorem 2 of [13], i.e., to establish the optimal decay rate $O(|\lambda|^{-1/2})$ for oscillatory integral operators with arbitrary analytic phases $S(x,y)$ and amplitudes with a damping factor $|S''_{xy}(x,y)|^{1/2}$. A paper containing these results is being prepared for publication.

(b) There is evidence that the above sharp statements for analytic phases may not be valid for arbitrary smooth phases. However, we expect that the sharp result will still hold in the C^∞-case under an additional *finite multiplicity* hypothesis: the ideal generated by S'''_{xxy} and S'''_{xyy} has finite codimension. We shall return to these matters elsewhere.

(c) Another set of questions deal with lower Newton decay rates for C^∞-phases. The estimate (1.5) with loss of ε for the case when condition (1.4) holds is in [15] (although not explicitly stated there). Since then, we have been informed by Seeger that he can also obtain results similar to ours, but with a loss of ε. In this connection, we should mention that A. Carbery has pointed out a flaw in the application of the stopping-time argument in our paper "On a stopping process for oscillatory integrals", *J. Geom. Anal.*, 4 (1994), 105–120. As a result, the logarithmic loss of decay claimed there is not established. We hope to return to this point in the future.

References

[1] ARNOLD, V., VARCHENKO, A. & GUSSEIN-ZADE, S., *Singularités des applications différentiables.* Mir, Moscow, 1986.

[2] CONNOR, J. L., CURTIS, P. R. & YOUNG, R. A. W., Uniform asymptotics of oscillating integrals: applications in chemical physics, in *Wave Asymptotics (Manchester,* 1990), pp. 24–42. Cambridge Univ. Press, Cambridge, 1992.

[3] GREENLEAF, A. & SEEGER, A., Fourier integral operators with fold singularities. *J. Reine Angew. Math.*, 445 (1994), 35–56.

[4] GREENLEAF, A. & UHLMANN, G., Composition of some singular Fourier integral operators and estimates for the X-ray transform, I. *Ann. Inst. Fourier*, 40 (1990), 443–466; II. *Duke Math. J.*, 64 (1991), 413–419.

[5] HÖRMANDER, L., Oscillatory integrals and multipliers on FL^p. *Ark. Mat.*, 11 (1973), 1–11.

[6] KENIG, C., PONCE, G. & VEGA, L., Oscillatory integrals and regularity of wave equations. *Indiana Math. J.*, 40 (1991), 33–69.

[7] MA, S., Forthcoming Ph.D. Thesis, Columbia University.

[8] MELROSE, R. & TAYLOR, M., Near peak scattering and the corrected Kirchhoff approximation for a convex obstacle. *Adv. in Math.*, 44 (1985), 242–315.

[9] PAN, Y. & SOGGE, C., Oscillatory integrals associated with canonical folding relations. *Colloq. Math.*, 40 (1990), 413–419.

[10] PHONG, D. H. & STEIN, E. M., Radon transforms and torsion. *Internat. Math. Res. Notices*, 4 (1991), 49–60.

[11] — Oscillatory integrals with polynomial phases. *Invent. Math.*, 110 (1992), 39–62.

[12] — Operator versions of the van der Corput lemma and Fourier integral operators. *Math. Res. Lett.*, 1 (1994), 27–33.

[13] — Models of degenerate Fourier integral operators and Radon transforms. *Ann. of Math.*, 140 (1994), 703–722.

[14] SAKS, S. & ZYGMUND, A., *Analytic Functions.* Elsevier, Amsterdam–London–New York, 1971.

[15] SEEGER, A., Degenerate Fourier integral operators in the plane. *Duke Math. J.*, 71 (1993), 685–745.

[16] SIEGEL, C. L., *Complex Function Theory,* Vol. I. Wiley-Interscience, New York, 1969.

[17] TAYLOR, M., Propagation, reflection, and diffraction of singularities of solutions to wave equations. *Bull. Amer. Math. Soc.*, 84 (1978), 589–611.

[18] VARCHENKO, A., Newton polyhedra and estimations of oscillatory integrals. *Functional Anal. Appl.*, 18 (1976), 175–196.

D. H. Phong
Department of Mathematics
Columbia University
New York, NY 10027
U.S.A.
phong@math.columbia.edu

E. M. Stein
Department of Mathematics
Princeton University
Princeton, NJ 08544
U.S.A.
stein@math.princeton.edu

Received June 28, 1996