ANALYSIS & PDE

Volume 16 No. 1 2023

msp

Analysis & PDE

msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2023 is US $405/year for the electronic version, and $630/year (+$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS AND PDE
Vol. 16 (2023), No. 1, pp. 1–34
DOI: 10.2140/apde.2023.16.1

VOLUME COMPARISON WITH RESPECT TO SCALAR CURVATURE

WEI YUAN

Dedicated to Nankai University
on its 100th Anniversary
(1919 - 2019)

We investigate the volume comparison with respect to scalar curvature. In particular, we show the volume comparison holds for small geodesic balls of metrics near a V-static metric. For closed manifolds, we prove the volume comparison for metrics near a strictly stable Einstein metric. As applications, we give a partial answer to a conjecture of Bray and recover a result of Besson, Courtois and Gallot, which partially confirms a conjecture of Schoen about closed hyperbolic manifolds. Applying analogous techniques, we obtain a different proof of a local rigidity result due to Dai, Wang and Wei, which shows it admits no metric with positive scalar curvature near strictly stable Ricci-flat metrics.

1. Introduction

The volume comparison theorem is a fundamental result in Riemannian geometry. It is a powerful tool in geometric analysis and frequently used in solving various problems.

The classic volume comparison theorem states that the volume of a complete manifold is upper bounded by the round sphere if its Ricci curvature is lower bounded by a corresponding positive constant. A natural question is whether we can replace the assumption on Ricci curvature by the one on scalar curvature.

In general, scalar curvature is not sufficient to control the volume. This is a straightforward consequence of a result by Corvino, Eichmair and Miao [Corvino et al. 2013]. In order to state it, we need the following fundamental concept, which was introduced in [Miao and Tam 2009].

Definition. Let $(M^n, \bar{g})$ be an n-dimensional Riemannian manifold. We say $\bar{g}$ is a V-static metric if there is a smooth function $f \not\equiv 0$ and a constant $\kappa \in \mathbb{R}$ that solve the *V-static equation*

$$\gamma_{\bar{g}}^* f = \nabla_{\bar{g}}^2 f - \bar{g}\Delta_{\bar{g}} f - f\,\mathrm{Ric}_{\bar{g}} = \kappa \bar{g}, \tag{1-1}$$

where $\gamma_{\bar{g}}^* : C^\infty(M) \to S_2(M)$ is the formal L^2-adjoint of $\gamma_{\bar{g}} := DR_{\bar{g}}$, the linearization of scalar curvature at $\bar{g}$. We also say a quadruple $(M, \bar{g}, f, \kappa)$ is a *V-static space*.

This work was supported by the NSFC (grants 12071489, 12025109, 11601531, and 11521101) .
MSC2010: primary 53C20; secondary 53C23, 53C24, 58J37.
Keywords: scalar curvature, volume comparison, V-static metric, stable Einstein metric, Bray's conjecture, Schoen's conjecture.

Remark 1.1. A fundamental property of a V-static metric is that its scalar curvature $R_{\bar{g}}$ is a constant for M connected; see Proposition 2.1 in [Corvino et al. 2013]. By taking the trace of (1-1), we can see that f satisfies the linear elliptic equation

$$\Delta_{\bar{g}} f + \frac{R_{\bar{g}}}{n-1} f + \frac{n\kappa}{n-1} = 0. \tag{1-2}$$

In particular, f is an eigenfunction for the Laplacian if $\kappa = 0$.

Einstein metrics are in particular V-static, which can be easily seen by taking the function f to be a constant. In this sense, we can view V-static metrics as a generalization of Einstein metrics. Another class of special V-static metrics are *vacuum static metrics* when we take $\kappa = 0$. They can be used to construct an important category of solutions to *Einstein field equations* in general relativity [Qing and Yuan 2013]. The classification of V-static spaces is a crucial problem in understanding the interplay between scalar curvature and volume. For more results, please refer to [Baltazar and Ribeiro 2017; Barros et al. 2015; Corvino et al. 2013; Miao and Tam 2009; 2012].

Now we state a deformation result associated with the concept of V-static metrics.

Theorem 1.2 (Corvino, Eichmair and Miao [Corvino et al. 2013]). *Let $(M^n, \bar{g})$ be a Riemannian manifold and $\Omega \subset M$ be a precompact domain with smooth boundary. Suppose $(\Omega, \bar{g})$ is not V-static, i.e., the V-static equation (1-1) only admits the trivial solution: $f \equiv 0$ and $\kappa = 0$ in $C^\infty(\Omega) \times \mathbb{R}$. Then for any Ω_0 compactly contained in Ω, there exists a constant $\delta_0 > 0$ such that for any $(\rho, V) \in C^\infty(M) \times \mathbb{R}$ with* $\operatorname{supp}(R_{\bar{g}} - \rho) \subset \Omega_0$ *and*

$$\|R_{\bar{g}} - \rho\|_{C^1(\Omega, \bar{g})} + |V_\Omega(\bar{g}) - V| < \delta_0,$$

there exists a metric g on M such that $\operatorname{supp}(g - \bar{g}) \subset \Omega,\ R_g = \rho$ *and* $V_\Omega(g) = V$.

This result suggests that for a non-V-static domain, the information of scalar curvature is not sufficient to give a volume comparison: we can take either $V > V_\Omega(\bar{g})$ or $V < V_\Omega(\bar{g})$, but with $\rho > R_{\bar{g}}$ in Ω. In either case, we can find a metric g realizing (ρ, V) on Ω and it shows that no volume comparison holds for non-V-static domains.

However, the volume comparison with respect to scalar curvature indeed holds for some special metrics. For instance, Miao and Tam [2012] proved a rigidity result for the upper hemisphere with respect to nondecreasing scalar curvature and volume. They also showed that a similar result holds for Euclidean domains. Note that since all space forms are V-static, it is natural to ask whether all V-static spaces admit such a volume comparison result.

Inspired by the rigidity of vacuum static metrics [Qing and Yuan 2016] and related work [Miao and Tam 2012], we obtain a volume comparison theorem with respect to scalar curvature for sufficiently small geodesic balls, if appropriate boundary conditions on induced metric $g|_{T\partial B_r(p)}$ and mean curvature H_g are posed.

Theorem A. *For $n \geq 3$, suppose $(M^n, \bar{g}, f, \kappa)$ is a V-static space. For any $p \in M$ with $f(p) > 0$, there exist positive constants r_0 and ε_0 such that for any geodesic ball $B_r(p) \subset M$ with radius $r \in (0, r_0)$ and metric g on $B_r(p)$ satisfying*

- $R_g \geq R_{\bar{g}}$ *in* $B_r(p)$,
- $H_g \geq H_{\bar{g}}$ *on* $\partial B_r(p)$,
- $g|_{T\partial B_r(p)} = \bar{g}|_{T\partial B_r(p)}$,
- $\|g - \bar{g}\|_{C^2(B_r(p),\bar{g})} < \varepsilon_0$,

the following volume comparison holds:

- *if* $\kappa < 0$, *then*
$$V_\Omega(g) \leq V_\Omega(\bar{g}),$$
- *if* $\kappa > 0$, *then*
$$V_\Omega(g) \geq V_\Omega(\bar{g}),$$

with equality holding in either case if and only if the metric g *is isometric to* $\bar{g}$.

Remark 1.3. If $f(p) < 0$, we only need to replace (f, κ) by $(-f, -\kappa)$, and the reversed volume comparison follows.

Remark 1.4. If $\kappa = 0$, then V-static metrics are in particular vacuum static, and hence g is isometric to $\bar{g}$ according to [Qing and Yuan 2016]. Thus Theorem A is an extension for the rigidity of vacuum static metrics.

In general, the function f may change its sign on a closed V-static manifold. For example, we can take $f := 1 + 2x_{n+1}$ on the unit sphere $\mathbb{S}^n$, where x_{n+1} is the height-function of $\mathbb{S}^n \hookrightarrow \mathbb{R}^{n+1}$. Hence the volume comparison may not hold in this case. However, for some special V-static spaces, the volume comparison with respect to scalar curvature might still hold for closed manifolds. Here and throughout this article, we say a manifold is *closed* if it is compact without boundary.

Schoen [1989] proposed a well-known conjecture that the Yamabe invariant of a given closed hyperbolic manifold is achieved by its canonical metric. This problem involves all possible metrics on a given hyperbolic manifold and it is obviously very difficult to solve. However, it can be shown that this conjecture is in fact equivalent to the following volume comparison problem.

Schoen's conjecture. *For* $n \geq 3$, *let* $(M^n, \bar{g})$ *be a closed hyperbolic manifold. Then for any metric* g *on* M *with*
$$R_g \geq R_{\bar{g}},$$
the volume comparison
$$V_M(g) \geq V_M(\bar{g})$$
holds.

The equivalence of the aforementioned Schoen's conjectures are known by experts. For the convenience of readers, we include a proof in the appendix.

Schoen's conjecture is known to hold for 3-manifolds due to works of Hamilton [1999] on nonsingular Ricci flow and Perelman [2002; 2003] on geometrization of 3-manifolds (also see [Agol et al. 2007] for a generalization). For higher dimensions, Besson, Courtois and Gallot [Besson et al. 1991] verified it for metrics C^2-close to the canonical metric. They also proved that the volume comparison holds without assuming g is close to $\bar{g}$ if one replaces the assumption on scalar curvature by Ricci curvature [Besson et al. 1995], which can be viewed as evidence that Schoen's conjecture holds for higher dimensions.

For the case of positive curvature, Bray proposed the following conjecture.

Bray's conjecture. *For $n \geq 3$, there is a positive constant $\varepsilon_n < 1$ such that for any complete manifold (M^n, g) with scalar curvature*

$$R_g \geq n(n-1)$$

and Ricci curvature

$$\mathrm{Ric}_g \geq \varepsilon_n (n-1) g,$$

the volume comparison

$$V_M(g) \leq V_{\mathbb{S}^n}(g_{\mathbb{S}^n})$$

holds, where $\mathbb{S}^n$ is the unit round sphere and $g_{\mathbb{S}^n}$ is the canonical metric.

Remark 1.5. Unlike Schoen's conjecture, there is an additional assumption on Ricci curvature in the positive curvature case. In fact, this assumption is necessary; see [Bray 1997] for details.

For this conjecture, Bray [1997] verified it for three dimensional manifolds and gave an estimate for ε_3. Later, Gursky and Viaclovsky [2004] showed that $\varepsilon_3 \leq \frac{1}{2}$, and Brendle [2012] proved the rigidity of volume comparison for $\varepsilon_3 = \frac{1}{2}$. For higher dimensions, Zhang [2019] gave a partial answer.

Before stating our result, we first recall the following well-known concept associated with an Einstein metric.

Definition 1.6 (stability of Einstein metrics). For $n \geq 3$, suppose $(M^n, \bar{g})$ is a closed Einstein manifold. The metric $\bar{g}$ is said to be *stable* if

$$\min \operatorname{spec}_{\mathrm{TT}}(-\Delta_E^{\bar{g}}) = \inf_{0 \not\equiv h \in S_2^{\mathrm{TT}}(M)} \frac{\int_M \langle h, -\Delta_E^{\bar{g}} h \rangle_{\bar{g}} \, dv_{\bar{g}}}{\int_M |h|_{\bar{g}}^2 \, dv_{\bar{g}}} \geq 0, \tag{1-3}$$

where $\Delta_E^{\bar{g}} := \Delta_{\bar{g}} + 2\,\mathrm{Rm}_{\bar{g}}$ is the *Einstein operator* and

$$S_{2,\bar{g}}^{\mathrm{TT}}(M) := \{h \in S_2(M) : \delta_{\bar{g}} h = 0,\ \mathrm{tr}_{\bar{g}}\, h = 0\}$$

is the space of *transverse-traceless symmetric* 2*-tensors* on $(M, \bar{g})$. Moreover, $\bar{g}$ is called *strictly stable* if the inequality in (1-3) is strict.

Stability is a crucial concept in the study of Einstein manifolds. There are several equivalent way to define it, we adopt the one involving the Einstein operator for our convenience. For more information, please refer to [Besse 1987; Dai et al. 2005; 2007; Kröncke 2013].

Our main result about volume comparison for Einstein manifolds is the following:

Theorem B. *Suppose $(M^n, \bar{g})$ is a closed strictly stable Einstein manifold with*

$$\mathrm{Ric}_{\bar{g}} = (n-1)\lambda \bar{g},$$

where $\lambda \neq 0$ is a constant. There exists a constant $\varepsilon_0 > 0$ such that for any metric g on M which satisfies

$$R_g \geq n(n-1)\lambda$$

and

$$\|g - \bar{g}\|_{C^2(M,\bar{g})} < \varepsilon_0,$$

the following volume comparison holds:

- *if* $\lambda > 0$, *then*
$$V_M(g) \leq V_M(\bar{g}),$$
- *if* $\lambda < 0$, *then*
$$V_M(g) \geq V_M(\bar{g}).$$

Moreover, *the equality holds in either case if and only if the metric g is isometric to $\bar{g}$.*

Remark 1.7. Suppose the reference metric $\bar{g}$ is Kähler–Einstein with negative scalar curvature and all infinitesimal complex deformations of its complex structure are integrable. Applying a delicate utilization of the functional

$$K(g) = \int_M |R_g|^{n/2} \, dv_g$$

and the Yamabe functional

$$Y(g) = \frac{\int_M R_g \, dv_g}{(V_M(g))^{(n-2)/n}},$$

Dai, Wang and Wei proved that the volume comparison with respect to scalar curvature holds for metrics near $\bar{g}$; see Theorem 1.5 in [Dai et al. 2007]. In fact, their result can be extended to strictly stable Einstein metrics with negative scalar curvature.

Remark 1.8. The above volume comparison does not hold for Ricci-flat metrics: by taking $g = c^2\bar{g}$ for a constant $c > 0$, we have the scalar curvature $R_g = R_{\bar{g}} = 0$, but the volume $V_M(g)$ can be either larger or smaller than $V_M(\bar{g})$ depending on whether $c > 1$ or $c < 1$.

Remark 1.9. The stability assumption in the theorem is necessary. Macbeth constructed an example of an Einstein manifold which shows the volume comparison fails if we lack stability (personal communication, 2019). See Proposition 5.9 for more details.

Remark 1.10. Our approach in fact works for other curvatures as well. Please see [Lin and Yuan 2022] for a volume comparison theorem of Q-curvature for strictly stable positive Einstein manifolds.

It is well known that hyperbolic metrics are strictly stable as special Einstein metrics and hence Theorem B provides a partial answer to Schoen's conjectures, which recovers the following result.

Corollary A (Besson, Courtois and Gallot [Besson et al. 1991]). *For $n \geq 3$, let $(M^n, \bar{g})$ be a closed hyperbolic manifold. There exists a constant $\varepsilon_0 > 0$ such that for any metric g on M with scalar curvature*

$$R_g \geq R_{\bar{g}}$$

and

$$\|g - \bar{g}\|_{C^2(M,\bar{g})} < \varepsilon_0,$$

we have

$$V_M(g) \geq V_M(\bar{g}),$$

where equality holds if and only if the metric g is isometric to $\bar{g}$.

Similarly, the spherical metric is also strictly stable (Example 3.1.2 in [Kröncke 2013]), and we obtain a partial answer to Bray's conjecture.

Corollary B. *For $n \geq 3$, let $(\mathbb{S}^n, g_{\mathbb{S}^n})$ be the unit round sphere. There exists a constant $\varepsilon_0 > 0$ such that for any metric g on $\mathbb{S}^n$ with scalar curvature*

$$R_g \geq n(n-1)$$

and

$$\|g - g_{\mathbb{S}^n}\|_{C^2(\mathbb{S}^n, g_{\mathbb{S}^n})} < \varepsilon_0,$$

we have

$$V_{\mathbb{S}^n}(g) \leq V_{\mathbb{S}^n}(g_{\mathbb{S}^n}),$$

where equality holds if and only if the metric g is isometric to $g_{\mathbb{S}^n}$.

Remark 1.11. For metrics close to the canonical spherical metric, the assumption on Ricci curvature in Bray's conjecture holds automatically.

Remark 1.12. Corvino, Eichmair and Miao constructed a metric on the upper hemisphere which satisfies the scalar comparison but has arbitrarily large volume; see Proposition 6.2 in [Corvino et al. 2013]. In fact, by gluing a lower hemisphere, we can get a metric on the whole sphere with scalar curvature no less than $n(n-1)$ but with larger volume.

In the research of scalar curvature, a fundamental question is whether a given manifold admits a metric of positive scalar curvature. A well-known result due to Schoen and Yau [1979a; 1979b] and Gromov and Lawson [1980; 1983] is the rigidity of tori, which states that there is no metric of positive scalar curvature on tori. For an excellent survey, please refer to [Rosenberg 2007].

In [Dai et al. 2005], Dai, Wang and Wei studied the existence of metrics with positive scalar curvature on a Riemannian manifold with nonzero parallel spinors. Through investigations of variational properties for the first eigenvalue of the conformal Laplacian, they proved the local rigidity of scalar curvature near the reference metric. Note that their proof can be applied to closed strictly stable Ricci-flat manifolds.

Applying techniques similar to the argument for Theorem B, we obtain the local rigidity of strictly stable Ricci-flat manifolds, which generalizes a result of Fischer and Marsden [1975] about local rigidity of tori with a different approach than in [Dai et al. 2005]:

Theorem C (Dai, Wang and Wei [Dai et al. 2005]). *Suppose $(M^n, \bar{g})$ is a strictly stable Ricci-flat manifold. Then there exists a constant $\varepsilon_0 > 0$ such that for any metric g on M satisfying*

$$R_g \geq 0$$

and

$$\|g - \bar{g}\|_{C^2(M, \bar{g})} < \varepsilon_0,$$

we have g is homothetic to $\bar{g}$. That is, we can find a constant $c > 0$ such that $g = c^2\bar{g}$. In particular, there is no metric with positive scalar curvature near $\bar{g}$.

Remark 1.13. Note that flat tori are merely stable, since the kernel of the Einstein operator is nontrivial and in fact

$$\dim\ker\Delta_E^{\bar{g}} \geq \frac{n(n+1)}{2} - 1.$$

It will be interesting to see whether there is an example of closed stable Ricci-flat manifold which admits a metric of positive scalar curvature near the reference metric.

Remark 1.14. Similar to Theorem B, our approach can also be applied to other curvatures. Please see [Lin and Yuan 2022] for an analogous result for Q-curvature.

The article is organized as follow: In Section 2, we collect notation and conventions used frequently in this article. In Section 3, we calculate some geometric variational formulas involved in the next two sections. In Section 4, we study the volume comparison for geodesic balls in V-static spaces. In Section 5, we investigate the volume comparison for non-Ricci-flat strictly stable Einstein manifolds and the rigidity phenomenon of strictly stable Ricci-flat manifolds. In the Appendix, we present a proof for equivalence of two conjectures proposed by Schoen.

2. Notation and conventions

In this section, we collect notation frequently used and conventions adopted in this article for the convenience of readers. Please note that *all calculations are performed in the reference metric $\bar{g}$.*

Let $(\Omega^n, \bar{g})$ be an n-dimensional compact manifold possibly with C^1-boundary $\Sigma := \partial\Omega$:

(1) Indices of coordinates components:
- Greek indices run through $1, \ldots, n$;
- Latin indices run through $1, \ldots, n-1$.

(2) Connections:
- connection on Ω with respect to $\bar{g}$: $\nabla_{\bar{g}}$;
- connection on Σ with respect to $\bar{g}|_{T\Sigma}$: ∇^{Σ}.

(3) Volume forms:
- volume form on Ω with respect to $\bar{g}$: $dv_{\bar{g}}$;
- volume form on Σ with respect to $\bar{g}|_{T\Sigma}$: $d\sigma_{\bar{g}}$.

(4) Curvatures:
- Riemann curvature tensor $\mathrm{Rm}_{\bar{g}}$: $R_{\alpha\beta\gamma\delta}$;
- Ricci curvature tensor $\mathrm{Ric}_{\bar{g}}$: $R_{\beta\gamma} = \bar{g}^{\alpha\delta}R_{\alpha\beta\gamma\delta}$;
- scalar curvature $R_{\bar{g}}$: $R_{\bar{g}} = \bar{g}^{\beta\gamma}R_{\beta\gamma}$;
- second fundamental form $A_{\bar{g}}$: $A^{\bar{g}}_{ij} = \frac{1}{2}\partial_{\nu_{\bar{g}}}\bar{g}_{ij}$;
- mean curvature $H_{\bar{g}}$: $H_{\bar{g}} = \bar{g}^{ij}A^{\bar{g}}_{ij}$.

(5) Spaces:

- space of all smooth Riemannian metrics on Ω: $\mathcal{M}_\Omega$;
- space of all smooth diffeomorphisms of Ω: $\mathscr{D}(\Omega)$;
- a local slice through the metric $\bar{g}$: $\mathcal{S}_{\bar{g}}$;
- symmetric 2-tensors on Ω: $S_2(\Omega)$;
- TT-tensors on $(\Omega, \bar{g})$: $S^{\mathrm{TT}}_{2,\bar{g}}(\Omega) = \{h \in S_2(\Omega) : \delta_{\bar{g}} h = 0,\ \mathrm{tr}_{\bar{g}}\, h = 0\}$.

(6) Operators:

- Multiplication and inner product of symmetric 2-tensors:

$$(h \times k)_{\alpha\delta} := \bar{g}^{\beta\gamma} h_{\alpha\beta} k_{\gamma\delta} \quad \text{and} \quad \langle h, k\rangle_{\bar{g}} = h \cdot k := \bar{g}^{\alpha\delta}(h \times k)_{\alpha\delta} = \bar{g}^{\alpha\delta}\bar{g}^{\beta\gamma} h_{\alpha\beta} k_{\gamma\delta}.$$

  In particular,

$$(h^2)_{\alpha\beta} = \bar{g}^{\gamma\delta} h_{\alpha\gamma} h_{\delta\beta} \quad \text{and} \quad \mathrm{Ric}_{\bar{g}} \cdot h := R_{\beta\gamma} h^{\beta\gamma}.$$

- Riemann curvature tensor as an operator on symmetric 2-tensors:

$$(\mathrm{Rm}_{\bar{g}} \cdot h)_{\beta\gamma} := R_{\alpha\beta\gamma\delta} h^{\alpha\delta} \quad \text{and} \quad \langle \mathrm{Rm}_{\bar{g}} \cdot h, h\rangle_{\bar{g}} := R_{\alpha\beta\gamma\delta} h^{\alpha\delta} h^{\beta\gamma}.$$

- A combination involving curvature:

$$\mathscr{R}_{\bar{g}}(h,h) := \langle \mathrm{Rm}_{\bar{g}} \cdot h, h\rangle_{\bar{g}} + 2(\mathrm{Ric}_{\bar{g}} \cdot h)(\mathrm{tr}_{\bar{g}}\, h) - \frac{2R_{\bar{g}}}{n-1}(\mathrm{tr}_{\bar{g}}\, h)^2.$$

- Formal L^2-adjoint of covariant differentiation:

$$\delta_{\bar{g}} = -\,\mathrm{div}_{\bar{g}}, \quad (\delta_{\bar{g}} h)_\beta = -\nabla^\alpha_{\bar{g}} h_{\alpha\beta}.$$

- Einstein operator:

$$\Delta^{\bar{g}}_E h = \Delta_{\bar{g}} h + 2\,\mathrm{Rm}_{\bar{g}} \cdot h.$$

- Linearization of scalar curvature:

$$\gamma_{\bar{g}} h = -\Delta_{\bar{g}}(\mathrm{tr}_{\bar{g}}\, h) + \delta^2_{\bar{g}} h - \mathrm{Ric}_{\bar{g}} \cdot h.$$

- Formal L^2-adjoint of $\gamma_{\bar{g}}$:

$$\gamma^*_{\bar{g}} f = \nabla^2_{\bar{g}} f - \bar{g}\Delta_{\bar{g}} f - f\,\mathrm{Ric}_{\bar{g}}\,.$$

3. Geometric variational formulas

In this section, we give variational formulas for geometric functionals involved later in the argument. Throughout this section, Ω is assumed to be a compact manifold possibly with C^1-boundary $\Sigma := \partial\Omega$. In the case $\Sigma \neq \varnothing$, let

$$\{e_1, \ldots, e_{n-1},\ e_n = \nu_{\bar{g}}\}$$

be an orthonormal frame on Σ such that the $\{e_i\}_{i=1}^{n-1}$ are tangent to Σ and $\nu_{\bar{g}}$ is the outward normal vector field of Σ with respect to the metric $\bar{g}$. We also denote the induced connection on Σ by ∇^Σ.

We begin with recalling well-known variational formulas of scalar curvature; for detailed calculations, please refer to [Fischer and Marsden 1975; Yuan 2015].

Lemma 3.1. *The first and second variations of scalar curvature are*

$$DR_{\bar{g}} \cdot h = -\Delta_{\bar{g}}(\operatorname{tr}_{\bar{g}} h) + \delta^2_{\bar{g}} h - \operatorname{Ric}_{\bar{g}} \cdot h, \tag{3-1}$$

and

$$D^2 R_{\bar{g}} \cdot (h, h) = -2\gamma_{\bar{g}}(h^2) - \Delta_{\bar{g}} |h|^2_{\bar{g}} - \tfrac{1}{2}|\nabla_{\bar{g}} h|^2_{\bar{g}} - \tfrac{1}{2}|d(\operatorname{tr}_{\bar{g}} h)|^2_{\bar{g}} + 2\langle h, \nabla^2_{\bar{g}}(\operatorname{tr}_{\bar{g}} h)\rangle_{\bar{g}} - 2\langle \delta_{\bar{g}} h, d(\operatorname{tr}_{\bar{g}} h)\rangle_{\bar{g}} + \nabla_\alpha h_{\beta\gamma} \nabla^\beta h^{\alpha\gamma} \tag{3-2}$$

for any $h \in S_2(\Omega)$.

For the mean curvature, its variations for the fixed induced boundary metric are given as follow, which was first shown in [Brendle and Marques 2011].

Lemma 3.2. *The first and second variations of mean curvature are*

$$DH_{\bar{g}} \cdot h = \tfrac{1}{2} h_{nn} H_{\bar{g}} - \nabla_i h_n{}^i + \tfrac{1}{2}\nabla_n h_i{}^i \tag{3-3}$$

and

$$D^2 H_{\bar{g}} \cdot (h, h) = \left(-\tfrac{1}{4} h^2_{nn} + \sum_{i=1}^{n-1} h^2_{in}\right) H_{\bar{g}} + h_{nn}\left(\nabla_i h_n{}^i - \tfrac{1}{2}\nabla_n h_i{}^i\right) \tag{3-4}$$

for any $h \in S_2(\Omega)$ with $h|_{T\partial\Omega} \equiv 0$.

For the volume functional, we provide a proof mainly based on a technique from linear algebra, which would be useful in calculating higher order variational formulas.

Lemma 3.3. *The first and second variations of volume are*

$$DV_{\Omega, \bar{g}} \cdot h = \frac{1}{2} \int_\Omega (\operatorname{tr}_{\bar{g}} h)\, dv_{\bar{g}} \tag{3-5}$$

and

$$D^2 V_{\Omega, \bar{g}} \cdot (h, h) = \frac{1}{4} \int_\Omega [(\operatorname{tr}_{\bar{g}} h)^2 - 2|h|^2_{\bar{g}}]\, dv_{\bar{g}} \tag{3-6}$$

for any $h \in S_2(\Omega)$.

Proof. Let A be an $n \times n$ symmetric matrix. Its characteristic polynomial is given by

$$\begin{aligned} p_A(\lambda) = \det(\lambda I - A) &= \sum_{k=0}^{n} (-1)^k \sigma_k(A) \lambda^{n-k} \\ &= \lambda^n - (\operatorname{tr} A)\lambda^{n-1} + \tfrac{1}{2}((\operatorname{tr} A)^2 - \operatorname{tr} A^2)\lambda^{n-2} + \sum_{k=3}^{n} (-1)^k \sigma_k(A)\lambda^{n-k}, \end{aligned}$$

where $\sigma_k(A)$ is the k-th elementary symmetric polynomial associated to the matrix A.

We choosing normal coordinates with respect to $\bar{g}$ centered at an interior point $x \in \Omega$, so that $\bar{g}_{\alpha\beta} = \delta_{\alpha\beta}$ at x. From the linear algebra fact mentioned above, we have the expansion

$$\det(\bar{g} + h) = 1 + (\operatorname{tr}_{\bar{g}} h) + \tfrac{1}{2}((\operatorname{tr}_{\bar{g}} h)^2 - |h|^2_{\bar{g}}) + O(|h|^3_{\bar{g}}),$$

and hence

$$\sqrt{\det(\bar{g}+h)} = 1 + \tfrac{1}{2}(\mathrm{tr}_{\bar{g}}\, h) + \tfrac{1}{8}((\mathrm{tr}_{\bar{g}}\, h)^2 - 2|h|_{\bar{g}}^2) + O(|h|_{\bar{g}}^3).$$

Immediately, this implies

$$DV_{\Omega,\bar{g}} \cdot h = \frac{1}{2}\int_\Omega (\mathrm{tr}_{\bar{g}}\, h)\, dv_{\bar{g}} \qquad \text{and} \qquad D^2V_{\Omega,\bar{g}} \cdot (h,h) = \frac{1}{4}\int_\Omega ((\mathrm{tr}_{\bar{g}}\, h)^2 - 2|h|_{\bar{g}}^2)\, dv_{\bar{g}},$$

respectively. □

In the rest of this section, we calculate variational formulas for some particularly designed functionals involving scalar curvature, mean curvature and volume.

Proposition 3.4. *For any $h \in S_2(\Omega)$ and $f \in C^\infty(\Omega)$,*

$$\int_\Omega (DR_{\bar{g}} \cdot h) f\, dv_{\bar{g}} = \int_\Omega \langle h, \gamma_{\bar{g}}^* f\rangle_{\bar{g}}\, dv_{\bar{g}} + \int_\Sigma [-(\partial_{\nu_{\bar{g}}}(\mathrm{tr}_{\bar{g}}\, h) + \langle \delta_{\bar{g}} h, \nu_{\bar{g}}\rangle_{\bar{g}}) f + (\mathrm{tr}_{\bar{g}}\, h)\partial_{\nu_{\bar{g}}} f - h(\nu_{\bar{g}}, \nabla_{\bar{g}} f)]\, d\sigma_{\bar{g}}.$$

Proof. It is straightforward to see that

$$\begin{aligned}\int_\Omega (DR_{\bar{g}} \cdot h) f\, dv_{\bar{g}} &= \int_\Omega (-\Delta_{\bar{g}}(\mathrm{tr}_{\bar{g}}\, h) + \delta_{\bar{g}}^2 h - \mathrm{Ric}_{\bar{g}} \cdot h) f\, dv_{\bar{g}} \\ &= \int_\Omega \langle h, \gamma_{\bar{g}}^* f\rangle_{\bar{g}}\, dv_{\bar{g}} + \int_\Sigma [-(\partial_{\nu_{\bar{g}}}(\mathrm{tr}_{\bar{g}}\, h) + \langle \delta_{\bar{g}} h, \nu_{\bar{g}}\rangle_{\bar{g}}) f + (\mathrm{tr}_{\bar{g}}\, h)\partial_{\nu_{\bar{g}}} f - h(\nu_{\bar{g}}, \nabla_{\bar{g}} f)]\, d\sigma_{\bar{g}},\end{aligned}$$

using Lemma 3.1 and integration by parts. □

Proposition 3.5. *For any $h \in S_2(\Omega)$ and $f \in C^\infty(\Omega)$,*

$$\begin{aligned}&\int_\Omega (D^2R_{\bar{g}} \cdot (h,h)) f\, dv_{\bar{g}} \\ &= \int_\Omega \big[-\tfrac{1}{2}|\nabla_{\bar{g}} h|_{\bar{g}}^2 - \tfrac{1}{2}|d(\mathrm{tr}_{\bar{g}}\, h)|_{\bar{g}}^2 + |\delta_{\bar{g}} h|^2 - 2\langle \delta_{\bar{g}} h, d(\mathrm{tr}_{\bar{g}}\, h)\rangle_{\bar{g}} + 2(\mathrm{tr}_{\bar{g}}\, h)(\delta_{\bar{g}}^2 h) + \mathscr{R}_{\bar{g}}(h,h)\big] f\, dv_{\bar{g}} \\ &\quad + \int_\Omega \Big[2(\mathrm{tr}_{\bar{g}}\, h)\Big(\langle h, \gamma_{\bar{g}}^* f\rangle_{\bar{g}} - 2\langle \delta_{\bar{g}} h, df\rangle_{\bar{g}} - \frac{1}{n-1}(\mathrm{tr}_{\bar{g}}\, h)(\mathrm{tr}_{\bar{g}}(\gamma_{\bar{g}}^* f))\Big) - 2\langle h, \delta_{\bar{g}} h \otimes df\rangle_{\bar{g}} - \langle \gamma_{\bar{g}}^* f, h^2\rangle_{\bar{g}}\Big]\, dv_{\bar{g}} \\ &\quad + \int_\Sigma [\partial_{\nu_{\bar{g}}}|h|_{\bar{g}}^2 + \langle \delta_{\bar{g}}(h^2), \nu_{\bar{g}}\rangle_{\bar{g}} + 2h(\nu_{\bar{g}}, \delta_{\bar{g}} h) + 2h(\nu_{\bar{g}}, \nabla_{\bar{g}}\, \mathrm{tr}_{\bar{g}}\, h) + 2(\mathrm{tr}_{\bar{g}}\, h)\langle \delta_{\bar{g}} h, \nu_{\bar{g}}\rangle_{\bar{g}}] f\, d\sigma_{\bar{g}} \\ &\quad + \int_\Sigma [h^2(\nu_{\bar{g}}, \nabla_{\bar{g}} f) - |h|_{\bar{g}}^2 \partial_{\nu_{\bar{g}}} f - 2(\mathrm{tr}_{\bar{g}}\, h) h(\nu_{\bar{g}}, \nabla_{\bar{g}} f)]\, d\sigma_{\bar{g}},\end{aligned}$$

where

$$\mathscr{R}_{\bar{g}}(h,h) := \langle \mathrm{Rm}_{\bar{g}} \cdot h, h\rangle_{\bar{g}} + 2(\mathrm{Ric}_{\bar{g}} \cdot h)(\mathrm{tr}_{\bar{g}}\, h) - \frac{2R_{\bar{g}}}{n-1}(\mathrm{tr}_{\bar{g}}\, h)^2.$$

Proof. By Lemma 3.1, we have

$$\begin{aligned}\int_\Omega (D^2R_{\bar{g}} \cdot (h,h)) f\, dv_{\bar{g}} &= \int_\Omega [-2\gamma_{\bar{g}}(h^2) - \Delta_{\bar{g}}|h|_{\bar{g}}^2 + 2\langle h, \nabla_{\bar{g}}^2(\mathrm{tr}_{\bar{g}}\, h)\rangle_{\bar{g}} + \nabla_\alpha h_{\beta\gamma}\nabla^\beta h^{\alpha\gamma}] f\, dv_{\bar{g}} \\ &\quad + \int_\Omega \big[-2\langle \delta_{\bar{g}} h, d(\mathrm{tr}_{\bar{g}}\, h)\rangle_{\bar{g}} - \tfrac{1}{2}|\nabla_{\bar{g}} h|_{\bar{g}}^2 - \tfrac{1}{2}|d(\mathrm{tr}_{\bar{g}}\, h)|_{\bar{g}}^2\big] f\, dv_{\bar{g}}.\end{aligned}$$

Integrating by parts,

$$
\begin{aligned}
-2&\int_\Omega (\gamma_{\bar g}(h^2)) f\, dv_{\bar g}\\
&= -2\int_\Omega \langle \gamma_{\bar g}^* f, h^2\rangle_{\bar g}\, dv_{\bar g} - 2\int_\Sigma [(\mathrm{tr}_{\bar g}(h^2))\partial_{\nu_{\bar g}} f - f\partial_{\nu_{\bar g}}(\mathrm{tr}_{\bar g}(h^2)) - h^2(\nu_{\bar g}, \nabla f) - \langle \delta_{\bar g}(h^2), \nu_{\bar g}\rangle_{\bar g} f]\, d\sigma_{\bar g}\\
&= -2\int_\Omega \langle \gamma_{\bar g}^* f, h^2\rangle_{\bar g}\, dv_{\bar g} + 2\int_\Sigma [(\partial_{\nu_{\bar g}}|h|^2_{\bar g} + \langle \delta_{\bar g}(h^2), \nu_{\bar g}\rangle_{\bar g}) f + h^2(\nu_{\bar g}, \nabla f) - |h|^2_{\bar g}\partial_{\nu_{\bar g}} f]\, d\sigma_{\bar g}
\end{aligned}
$$

and

$$
-\int_\Omega (\Delta_{\bar g}|h|^2) f\, dv_{\bar g} = -\int_\Omega (|h|^2\Delta_{\bar g} f)\, dv_{\bar g} - \int_\Sigma [f\partial_{\nu_{\bar g}}|h|^2_{\bar g} - |h|^2_{\bar g}\partial_{\nu_{\bar g}} f]\, d\sigma_{\bar g}.
$$

Also,

$$
\begin{aligned}
2&\int_\Omega \langle h, \nabla^2_{\bar g}(\mathrm{tr}_{\bar g} h)\rangle_{\bar g} f\, dv_{\bar g}\\
&= 2\int_\Omega [\langle \delta_{\bar g} h, d(\mathrm{tr}_{\bar g} h)\rangle f - \langle h, d(\mathrm{tr}_{\bar g} h)\otimes df\rangle_{\bar g}]\, dv_{\bar g} + 2\int_\Sigma h(\nu_{\bar g}, \nabla_{\bar g}(\mathrm{tr}_{\bar g} h)) f\, d\sigma_{\bar g}\\
&= 2\int_\Omega (\mathrm{tr}_{\bar g} h)[(\delta^2_{\bar g} h) f - 2\langle \delta_{\bar g} h, df\rangle_{\bar g} + \langle h, \nabla^2_{\bar g} f\rangle_{\bar g}]\, dv_{\bar g}\\
&\qquad + 2\int_\Sigma [(h(\nu_{\bar g}, \nabla_{\bar g}(\mathrm{tr}_{\bar g} h)) + (\mathrm{tr}_{\bar g} h)\langle \delta_{\bar g} h, \nu_{\bar g}\rangle_{\bar g}) f - (\mathrm{tr}_{\bar g} h) h(\nu_{\bar g}, \nabla_{\bar g} f)]\, d\sigma_{\bar g}\\
&= 2\int_\Omega (\mathrm{tr}_{\bar g} h)[(\delta^2_{\bar g} h) f - 2\langle \delta_{\bar g} h, df\rangle_{\bar g} + \langle h, \gamma^*_{\bar g} f\rangle_{\bar g} + (\mathrm{tr}_{\bar g} h)\Delta_{\bar g} f + (\mathrm{Ric}_{\bar g}\cdot h) f]\, dv_{\bar g}\\
&\qquad + 2\int_\Sigma [(h(\nu_{\bar g}, \nabla_{\bar g}(\mathrm{tr}_{\bar g} h)) + (\mathrm{tr}_{\bar g} h)\langle \delta_{\bar g} h, \nu_{\bar g}\rangle_{\bar g}) f - (\mathrm{tr}_{\bar g} h) h(\nu_{\bar g}, \nabla_{\bar g} f)]\, d\sigma_{\bar g}
\end{aligned}
$$

and

$$
\begin{aligned}
&\int_\Omega [\nabla_\alpha h_{\beta\gamma}\nabla^\beta h^{\alpha\gamma}] f\, dv_{\bar g}\\
&= -\int_\Omega h_\gamma{}^\beta [\nabla_\alpha\nabla_\beta h^{\alpha\gamma} f + \nabla_\beta h^{\alpha\gamma}\nabla_\alpha f]\, dv_{\bar g} + \int_\Sigma [h_{\beta\gamma}\nu_{\bar g\alpha}\nabla^\beta h^{\alpha\gamma}] f\, d\sigma_{\bar g}\\
&= -\int_\Omega h_\gamma{}^\beta [(\nabla_\beta\nabla_\alpha h^{\alpha\gamma} + R_{\alpha\beta\delta}{}^\alpha h^{\delta\gamma} + R_{\alpha\beta\delta}{}^\gamma h^{\alpha\delta}) f + \nabla_\beta h^{\alpha\gamma}\nabla_\alpha f]\, dv_{\bar g} + \int_\Sigma [h_{\beta\gamma}\nu_{\bar g\alpha}\nabla^\beta h^{\alpha\gamma}] f\, d\sigma_{\bar g}\\
&= -\int_\Omega [-\nabla_\beta h_\gamma{}^\beta\nabla_\alpha h^{\alpha\gamma} f - 2h_\gamma{}^\beta\nabla_\alpha h^{\alpha\gamma}\nabla_\beta f - h_\gamma{}^\beta h^{\alpha\gamma}\nabla_\beta\nabla_\alpha f + (\langle \mathrm{Ric}_{\bar g}, h^2\rangle_{\bar g} - \langle \mathrm{Rm}_{\bar g}\cdot h, h\rangle_{\bar g}) f]\, dv_{\bar g}\\
&\qquad + \int_\Sigma [(h_{\beta\gamma}\nu_{\bar g\alpha}\nabla^\beta h^{\alpha\gamma} - h^\beta_\gamma\nu_{\bar g\beta}\nabla_\alpha h^{\alpha\gamma}) f - h^\beta_\gamma h^{\alpha\gamma}\nu_{\bar g\beta}\nabla_\alpha f]\, d\sigma_{\bar g}\\
&= \int_\Omega [|\delta_{\bar g} h|^2_{\bar g} f - 2\langle h, \delta_{\bar g} h\otimes df\rangle_{\bar g} + \langle \nabla^2_{\bar g} f - f\,\mathrm{Ric}_{\bar g}, h^2\rangle_{\bar g} + \langle \mathrm{Rm}_{\bar g}\cdot h, h\rangle_{\bar g} f]\, dv_{\bar g}\\
&\qquad - \int_\Sigma [(\langle \delta_{\bar g}(h^2), \nu_{\bar g}\rangle_{\bar g} - 2h(\nu_{\bar g}, \delta_{\bar g} h)) f + h^2(\nu_{\bar g}, \nabla_{\bar g} f)]\, d\sigma_{\bar g}\\
&= \int_\Omega [|\delta_{\bar g} h|^2_{\bar g} f - 2\langle h, \delta_{\bar g} h\otimes df\rangle_{\bar g} + \langle \gamma^*_{\bar g} f + \bar g\Delta_{\bar g} f, h^2\rangle_{\bar g} + \langle \mathrm{Rm}_{\bar g}\cdot h, h\rangle_{\bar g} f]\, dv_{\bar g}\\
&\qquad - \int_\Sigma [(\langle \delta_{\bar g}(h^2), \nu_{\bar g}\rangle_{\bar g} - 2h(\nu_{\bar g}, \delta_{\bar g} h)) f + h^2(\nu_{\bar g}, \nabla_{\bar g} f)]\, d\sigma_{\bar g}.
\end{aligned}
$$

Combining the calculations above, we obtain

$$\begin{aligned}
&\int_\Omega (D^2R_{\bar g}\cdot(h,h))f\,dv_{\bar g}\\
&=\int_\Omega\big[-\tfrac12|\nabla_{\bar g}h|^2_{\bar g}-\tfrac12|d(\mathrm{tr}_{\bar g}h)|^2_{\bar g}+|\delta_{\bar g}h|^2_{\bar g}-2\langle\delta_{\bar g}h,d(\mathrm{tr}_{\bar g}h)\rangle_{\bar g}+\langle\mathrm{Rm}_{\bar g}\cdot h,h\rangle_{\bar g}+2(\mathrm{tr}_{\bar g}h)(\mathrm{Ric}_{\bar g}\cdot h)\big]f\,dv_{\bar g}\\
&\quad+\int_\Omega[2(\mathrm{tr}_{\bar g}h)((\delta^2_{\bar g}h)f+\langle h,\gamma^*_{\bar g}f\rangle_{\bar g}-2\langle\delta_{\bar g}h,df\rangle_{\bar g}+(\mathrm{tr}_{\bar g}h)\Delta_{\bar g}f)-2\langle h,\delta_{\bar g}h\otimes df\rangle_{\bar g}-\langle\gamma^*_{\bar g}f,h^2\rangle_{\bar g}]\,dv_{\bar g}\\
&\quad\quad+\int_\Sigma[(\partial_{\nu_{\bar g}}|h|^2_{\bar g}+\langle\delta_{\bar g}(h^2),\nu_{\bar g}\rangle_{\bar g}+2h(\nu_{\bar g},\delta_{\bar g}h))f-|h|^2_{\bar g}\partial_{\nu_{\bar g}}f+h^2(\nu_{\bar g},\nabla_{\bar g}f)]\,d\sigma_{\bar g}\\
&\quad\quad\quad+2\int_\Sigma[(h(\nu_{\bar g},\nabla_{\bar g}(\mathrm{tr}_{\bar g}h))+(\mathrm{tr}_{\bar g}h)\langle\delta_{\bar g}h,\nu_{\bar g}\rangle_{\bar g})f-(\mathrm{tr}_{\bar g}h)h(\nu_{\bar g},\nabla_{\bar g}f)]\,d\sigma_{\bar g}\\
&=\int_\Omega\big[-\tfrac12|\nabla_{\bar g}h|^2_{\bar g}-\tfrac12|d(\mathrm{tr}_{\bar g}h)|^2_{\bar g}+|\delta_{\bar g}h|^2_{\bar g}-2\langle\delta_{\bar g}h,d(\mathrm{tr}_{\bar g}h)\rangle_{\bar g}+2(\mathrm{tr}_{\bar g}h)(\delta^2_{\bar g}h)+\mathscr{R}_{\bar g}(h,h)\big]f\,dv_{\bar g}\\
&\quad+\int_\Omega\Big[2(\mathrm{tr}_{\bar g}h)\Big(\langle h,\gamma^*_{\bar g}f\rangle_{\bar g}-2\langle\delta_{\bar g}h,df\rangle_{\bar g}-\frac{1}{n-1}(\mathrm{tr}_{\bar g}h)(\mathrm{tr}_{\bar g}(\gamma^*_{\bar g}f))\Big)-2\langle h,\delta_{\bar g}h\otimes df\rangle_{\bar g}-\langle\gamma^*_{\bar g}f,h^2\rangle_{\bar g}\Big]\,dv_{\bar g}\\
&\quad\quad+\int_\Sigma[\partial_{\nu_{\bar g}}|h|^2_{\bar g}+\langle\delta_{\bar g}(h^2),\nu_{\bar g}\rangle_{\bar g}+2h(\nu_{\bar g},\delta_{\bar g}h)+2h(\nu_{\bar g},\nabla_{\bar g}(\mathrm{tr}_{\bar g}h))+2(\mathrm{tr}_{\bar g}h)\langle\delta_{\bar g}h,\nu_{\bar g}\rangle_{\bar g}]f\,d\sigma_{\bar g}\\
&\quad\quad\quad+\int_\Sigma[h^2(\nu_{\bar g},\nabla_{\bar g}f)-|h|^2_{\bar g}\partial_{\nu_{\bar g}}f-2(\mathrm{tr}_{\bar g}h)h(\nu_{\bar g},\nabla_{\bar g}f)]\,d\sigma_{\bar g},
\end{aligned}$$

where we used the fact that

$$\mathrm{tr}_{\bar g}(\gamma^*_{\bar g}f)=-(n-1)\Big(\Delta_{\bar g}f+\frac{R_{\bar g}}{n-1}f\Big)$$

and

$$\mathscr{R}_{\bar g}(h,h)=\langle\mathrm{Rm}_{\bar g}\cdot h,h\rangle_{\bar g}+2(\mathrm{Ric}_{\bar g}\cdot h)(\mathrm{tr}_{\bar g}h)-\frac{2R_{\bar g}}{n-1}(\mathrm{tr}_{\bar g}h)^2.$$ □

In particular, for V-static metrics we have the following identity.

Corollary 3.6. *Suppose $(\Omega,\bar g,f,\kappa)$ is a V-static space. Then for any $h\in\ker\delta_{\bar g}$ with $h|_{T\Sigma}\equiv 0$,*

$$\begin{aligned}
\int_\Omega(D^2R_{\bar g}\cdot(h,h))f\,dv_{\bar g}=&-\frac12\int_\Omega\Big[(|\nabla_{\bar g}h|^2_{\bar g}+|d(\mathrm{tr}_{\bar g}h)|^2_{\bar g}-2\mathscr{R}_{\bar g}(h,h))f+2\kappa\Big(|h|^2_{\bar g}+\frac{2}{n-1}(\mathrm{tr}_{\bar g}h)^2\Big)\Big]\,dv_{\bar g}\\
&-\int_\Sigma\Big[A^{ij}_{\bar g}h_{in}h_{jn}-\Big(h^2_{nn}-3\sum_{i=1}^{n-1}h^2_{in}\Big)H_{\bar g}+4h_{nn}\Big(\nabla_ih_n{}^i-\frac12\nabla_nh_i{}^i\Big)\Big]f\,d\sigma_{\bar g}\\
&-\int_\Sigma\Big[\Big(2h^2_{nn}+\sum_{i=1}^{n-1}h^2_{in}\Big)\partial_nf+2h_{nn}\sum_{i=1}^{n-1}h_{in}\partial_if\Big]\,d\sigma_{\bar g}.
\end{aligned}$$

Proof. Applying Proposition 3.5 with our assumptions,

$$\begin{aligned}
\int_\Omega(D^2R_{\bar g}\cdot(h,h))f\,dv_{\bar g}=&-\frac12\int_\Omega\Big[(|\nabla_{\bar g}h|^2_{\bar g}+|d(\mathrm{tr}_{\bar g}h)|^2_{\bar g}-2\mathscr{R}_{\bar g}(h,h))f+2\kappa\Big(|h|^2_{\bar g}+\frac{2}{n-1}(\mathrm{tr}_{\bar g}h)^2\Big)\Big]\,dv_{\bar g}\\
&+\int_\Sigma\big[(\partial_{\nu_{\bar g}}|h|^2_{\bar g}+\langle\delta_{\bar g}(h^2),\nu_{\bar g}\rangle_{\bar g}+2h(\nu_{\bar g},\nabla_{\bar g}(\mathrm{tr}_{\bar g}h)))f+h^2(\nu_{\bar g},\nabla_{\bar g}f)\\
&\qquad-|h|^2_{\bar g}\partial_{\nu_{\bar g}}f-2(\mathrm{tr}_{\bar g}h)h(\nu_{\bar g},\nabla_{\bar g}f)\big]\,d\sigma_{\bar g}.
\end{aligned}$$

For the boundary integral, we will rewrite it in terms of the orthonormal frame chosen for the boundary. Note that the identities

$$\Gamma^n_{ij} = -A^{\bar g}_{ij}, \qquad \Gamma^k_{jn} = A^k_j, \qquad \Gamma^i_{in} = H_{\bar g} \tag{3-7}$$

hold on Σ. Since

$$\delta_{\bar g} h = 0 \quad \text{and} \quad h_{ij} = 0, \quad i, j = 1, \dots, n-1,$$

we have

$$\langle \delta_{\bar g}(h^2), \nu_{\bar g}\rangle_{\bar g} = (\delta_{\bar g}(h^2))_n = -\nabla_\alpha (h_\beta{}^\alpha h_n{}^\beta) = -h_\beta{}^\alpha \nabla_\alpha h_n{}^\beta = -h_{nn}\nabla_n h_{nn} - h_n{}^i \nabla_i h_{nn} - h_n{}^i \nabla_n h_{in},$$

$$\partial_{\nu_{\bar g}} |h|^2_{\bar g} = \nabla_n |h|^2_{\bar g} = 2h_{nn}\nabla_n h_{nn} + 4h_n{}^i \nabla_n h_{in}$$

on Σ. Thus,

$$\begin{aligned}
&\partial_{\nu_{\bar g}} |h|^2_{\bar g} + \langle \delta_{\bar g}(h^2), \nu_{\bar g}\rangle_{\bar g} + 2h(\nu_{\bar g}, \nabla_{\bar g}(\operatorname{tr}_{\bar g} h)) \\
&\qquad = h_{nn}\nabla_n h_{nn} + 3h_n{}^i \nabla_n h_{in} - h_n{}^i \nabla_i h_{nn} + 2h_{nn}\nabla_n(\operatorname{tr}_{\bar g} h) + 2h_n{}^i \nabla_i(\operatorname{tr}_{\bar g} h) \\
&\qquad = 3h_{nn}\nabla_n h_{nn} + 3h_n{}^i \nabla_n h_{in} - h_n{}^i \nabla_i h_{nn} + 2h_{nn}\nabla_n h_i{}^i + 2h_n{}^i \nabla^\Sigma_i h_{nn} \\
&\qquad = -3h_{nn}\nabla_i h_n{}^i - 3h_n{}^i \nabla_j h_i{}^j - h_n{}^i \nabla_i h_{nn} + 2h_{nn}\nabla_n h_i{}^i + 2h_n{}^i \nabla^\Sigma_i h_{nn},
\end{aligned}$$

where we used the fact that

$$\nabla_n h_{n\alpha} = -(\delta_{\bar g} h)_\alpha - \nabla_i h_\alpha{}^i = -\nabla_i h_\alpha{}^i.$$

Moreover, from

$$\nabla_j h_i{}^j = \partial_j h_i{}^j + \Gamma^j_{j\alpha} h_i{}^\alpha - \Gamma^\alpha_{ji} h_\alpha{}^j = A^{\bar g}_{ij} h_n{}^j + H_{\bar g} h_{in}$$

and

$$\nabla_i h_{nn} = \partial_i h_{nn} - 2\Gamma^\alpha_{in} h_{\alpha n} = \nabla^\Sigma_i h_{nn} - 2A^{\bar g}_{ij} h_n{}^j,$$

we obtain

$$\begin{aligned}
&\partial_{\nu_{\bar g}} |h|^2_{\bar g} + \langle \delta_{\bar g}(h^2), \nu_{\bar g}\rangle_{\bar g} + 2h(\nu_{\bar g}, \nabla_{\bar g}(\operatorname{tr}_{\bar g} h)) \\
&\qquad = -A^{ij}_{\bar g} h_{in} h_{jn} - 3H_{\bar g} \sum_{i=1}^{n-1} h^2_{in} + h_n{}^i \nabla^\Sigma_i h_{nn} - 3h_{nn}\nabla_i h_n{}^i + 2h_{nn}\nabla_n h_i{}^i.
\end{aligned}$$

On the other hand,

$$h^2(\nu_{\bar g}, \nabla_{\bar g} f) - |h|^2_{\bar g} \partial_{\nu_{\bar g}} f - 2(\operatorname{tr}_{\bar g} h) h(\nu_{\bar g}, \nabla_{\bar g} f) = -\biggl(2h^2_{nn} + \sum_{i=1}^{n-1} h^2_{in}\biggr)\partial_n f - h_{nn} \sum_{i=1}^{n-1} h_{in}\partial_i f.$$

Integrating by parts,

$$\begin{aligned}
&\int_\Sigma [(\partial_{\nu_{\bar g}} |h|^2_{\bar g} + \langle \delta_{\bar g}(h^2), \nu_{\bar g}\rangle_{\bar g} + 2h(\nu_{\bar g}, \nabla_{\bar g}(\operatorname{tr}_{\bar g} h))) f + h^2(\nu_{\bar g}, \nabla_{\bar g} f) - |h|^2_{\bar g}\partial_{\nu_{\bar g}} f - 2(\operatorname{tr}_{\bar g} h) h(\nu_{\bar g}, \nabla_{\bar g} f)]\, d\sigma_{\bar g} \\
&\quad = -\int_\Sigma \biggl[\biggl(A^{ij}_{\bar g} h_{in} h_{jn} + 3H_{\bar g} \sum_{i=1}^{n-1} h^2_{in}\biggr) f + \biggl(2h^2_{nn} + \sum_{i=1}^{n-1} h^2_{in}\biggr)\partial_n f + 2h_{nn} \sum_{i=1}^{n-1} h_{in}\partial_i f\biggr]\, d\sigma_{\bar g} \\
&\qquad + \int_\Sigma (-h_{nn}\nabla^\Sigma_i h_n{}^i - 3h_{nn}\nabla_i h_n{}^i + 2h_{nn}\nabla_n h_i{}^i) f \, d\sigma_{\bar g}.
\end{aligned}$$

Note that

$$\begin{aligned}\nabla_i h_n{}^i &= \partial_i h_n{}^i + \Gamma^i_{i\alpha} h_n{}^\alpha - \Gamma^\alpha_{in} h_\alpha{}^i \\ &= \nabla^\Sigma_i h_n{}^i + H_{\bar g} h_{nn},\end{aligned}$$

and hence

$$\begin{aligned}\int_\Sigma [(\partial_{\nu_{\bar g}} |h|^2_{\bar g} + \langle \delta_{\bar g}(h^2), \nu_{\bar g}\rangle_{\bar g} + 2h(\nu_{\bar g}, \nabla_{\bar g}(\mathrm{tr}_{\bar g} h))) f + h^2(\nu_{\bar g}, \nabla_{\bar g} f) - |h|^2_{\bar g} \partial_{\nu_{\bar g}} f - 2(\mathrm{tr}_{\bar g} h) h(\nu_{\bar g}, \nabla_{\bar g} f)]\, d\sigma_{\bar g} \\ = -\int_\Sigma \left[A^{ij}_{\bar g} h_{in} h_{jn} - \left(h^2_{nn} - 3\sum_{i=1}^{n-1} h^2_{in} \right) H_{\bar g} + 4h_{nn} \left(\nabla_i h_n{}^i - \tfrac{1}{2} \nabla_n h_i{}^i \right) \right] f\, d\sigma_{\bar g} \\ - \int_\Sigma \left[\left(2h^2_{nn} + \sum_{i=1}^{n-1} h^2_{in} \right) \partial_n f + 2h_{nn} \sum_{i=1}^{n-1} h_{in} \partial_i f \right] d\sigma_{\bar g}. \quad \square\end{aligned}$$

In particular, for a special class of V-static spaces we have the following.

Corollary 3.7. *Suppose $(M^n, \bar g)$ is a closed Einstein manifold with*

$$\mathrm{Ric}_{\bar g} = (n-1)\lambda \bar g.$$

Then for any $h \in S^{\mathrm{TT}}_{2,\bar g}(M) \oplus (C^\infty(M) \cdot \bar g)$ we have

$$\int_M (D^2 R_{\bar g} \cdot (h, h))\, dv_{\bar g} = -\frac{1}{2} \int_M \left(-\langle h, \Delta^{\bar g}_E h\rangle_{\bar g} + \frac{n^2-2}{n^2} |d(\mathrm{tr}_{\bar g} h)|^2_{\bar g} - 2(n-1)\lambda |h|^2_{\bar g} \right) dv_{\bar g}.$$

Proof. According to the V-static equation (1-1), it is obvious that the Einstein manifold $(M^n, \bar g)$ is a V-static space with $f \equiv 1$ on M and $\kappa = -(n-1)\lambda$. By Corollary 3.6 we obtain

$$\int_M (D^2 R_{\bar g} \cdot (h, h))\, dv_{\bar g} = \int_M \left[-\tfrac{1}{2} |\nabla_{\bar g} h|^2_{\bar g} - \tfrac{1}{2} |d(\mathrm{tr}_{\bar g} h)|^2_{\bar g} + |\delta_{\bar g} h|^2_{\bar g} + \mathscr{R}_{\bar g}(h, h) + 2\lambda (\mathrm{tr}_{\bar g} h)^2 + (n-1)\lambda |h|^2_{\bar g} \right] dv_{\bar g}.$$

From our assumption,

$$\delta_{\bar g} h = -\frac{1}{n} d(\mathrm{tr}_{\bar g} h),$$

and hence

$$\int_M (D^2 R_{\bar g} \cdot (h, h))\, dv_{\bar g} = \int_M \left[-\tfrac{1}{2} |\nabla_{\bar g} h|^2_{\bar g} - \frac{n^2-2}{2n^2} |d(\mathrm{tr}_{\bar g} h)|^2_{\bar g} + \mathscr{R}_{\bar g}(h, h) + 2\lambda (\mathrm{tr}_{\bar g} h)^2 + (n-1)\lambda |h|^2_{\bar g} \right] dv_{\bar g}.$$

Since

$$\begin{aligned}\mathscr{R}_{\bar g}(h, h) &= \langle \mathrm{Rm}_{\bar g} \cdot h, h\rangle_{\bar g} + 2(\mathrm{Ric}_{\bar g} \cdot h)(\mathrm{tr}_{\bar g} h) - \frac{2R_{\bar g}}{n-1} (\mathrm{tr}_{\bar g} h)^2 \\ &= \langle \mathrm{Rm}_{\bar g} \cdot h, h\rangle_{\bar g} - 2\lambda (\mathrm{tr}_{\bar g} h)^2,\end{aligned}$$

we have

$$\int_M (D^2 R_{\bar g} \cdot (h, h))\, dv_{\bar g} = -\frac{1}{2} \int_M \left(-\langle h, \Delta^{\bar g}_E h\rangle_{\bar g} + \frac{n^2-2}{n^2} |d(\mathrm{tr}_{\bar g} h)|^2_{\bar g} - 2(n-1)\lambda |h|^2_{\bar g} \right) dv_{\bar g}. \quad \square$$

4. Volume comparison for V-static spaces

In this section, we will investigate the volume comparison for geodesic balls in generic V-static spaces.

Let Ω be an n-dimensional compact domain in a V-static space $(M^n, \bar{g}, f, \kappa)$ with C^1-boundary $\Sigma := \partial\Omega$. We define the functional

$$\mathscr{F}_{\Omega,\bar{g}}[g] := \int_\Omega R(g) f \, dv_{\bar{g}} + 2\int_\Sigma H(g) f \, d\sigma_{\bar{g}} - 2\kappa V_\Omega(g), \tag{4-1}$$

where

$$g \in \mathcal{M}_{\Omega,\Sigma,\bar{g}} := \{g \in \mathcal{M}_\Omega : g|_{T\Sigma} = \bar{g}|_{T\Sigma}\}$$

is a Riemannian metric on Ω that induces the same metric as $\bar{g}$ on the boundary Σ.

This functional is particularly designed for a given V-static space. The information of both volume and curvature is encoded in this single functional. It has excellent variational properties.

Proposition 4.1. *The V-static metric $\bar{g}$ is a critical point of the functional $\mathscr{F}_{\Omega,\bar{g}}[g]$. That is,*

$$D\mathscr{F}_{\Omega,\bar{g}} \cdot h = 0 \tag{4-2}$$

for any $h \in S_2(\Omega)$ with $h|_{T\partial\Omega} \equiv 0$.

Proof. Applying Proposition 3.4 together with Lemmas 3.2 and 3.3,

$$\begin{aligned} D\mathscr{F}_{\Omega,\bar{g}} \cdot h &= \int_\Omega (DR_{\bar{g}} \cdot h) f \, dv_{\bar{g}} + 2\int_{\partial\Omega} (DH_{\bar{g}} \cdot h) f \, d\sigma_{\bar{g}} - 2\kappa (DV_{\Omega,\bar{g}} \cdot h) \\ &= \int_\Omega [\langle h, \gamma_{\bar{g}}^* f\rangle_{\bar{g}} - \kappa(\operatorname{tr}_{\bar{g}} h)] \, dv_{\bar{g}} \\ &\qquad + \int_{\partial\Omega} [-(\partial_n(\operatorname{tr}_{\bar{g}} h) + (\delta_{\bar{g}} h)_n + 2\nabla_i h_n{}^i - \nabla_n h_i{}^i - h_{nn} H_{\bar{g}}) f - h_n{}^i \partial_i f] \, d\sigma_{\bar{g}}, \end{aligned}$$

where we used that $\operatorname{tr}_{\bar{g}} h = h_{nn}$ on $\partial\Omega$. Since

$$\nabla_i h_n{}^i = \partial_i h_n{}^i + \Gamma^i_{i\alpha} h_n{}^\alpha - \Gamma^\alpha_{in} h_\alpha{}^i = \nabla_i^\Sigma h_n{}^i + H_{\bar{g}} h_{nn},$$

we have

$$(\delta_{\bar{g}} h)_n = -\nabla_\alpha h_n{}^\alpha = -\nabla_i^\Sigma h_n{}^i - \nabla_n h_{nn} - H_{\bar{g}} h_{nn}.$$

Therefore

$$D\mathscr{F}_{\Omega,\bar{g}} \cdot h = \int_\Omega \langle h, \gamma_{\bar{g}}^* f - \kappa\bar{g}\rangle_{\bar{g}} \, dv_{\bar{g}} - \int_{\partial\Omega} [(\nabla_i^\Sigma h_n{}^i) f + h_n{}^i \partial_i f] \, d\sigma_{\bar{g}} = -\int_{\partial\Omega} \nabla_i^\Sigma (h_n{}^i f) \, d\sigma_{\bar{g}} = 0,$$

i.e., $\bar{g}$ is a critical point of $\mathscr{F}_{\Omega,\bar{g}}[g]$. □

The second variation that follows is a straightforward application of Lemmas 3.2 and 3.3 together with Corollary 3.6.

Proposition 4.2. *For any $h \in \ker \delta_{\bar g}$ with $h|_{T\Sigma} \equiv 0$, we have*

$$D^2\mathscr{F}_{\Omega,\bar g}\cdot(h,h) = -\frac{1}{2}\int_\Omega\Big[(|\nabla_{\bar g}h|^2_{\bar g} + |d(\mathrm{tr}_{\bar g}h)|^2_{\bar g} - 2\mathscr{R}_{\bar g}(h,h))f + \frac{n+3}{n-1}(\mathrm{tr}_{\bar g}h)^2\kappa\Big]dv_{\bar g}$$
$$-\int_\Sigma\Big[\Big(A^{ij}_{\bar g}h_{in}h_{jn} - \frac{1}{2}\Big(h^2_{nn} - 2\sum_{i=1}^{n-1}h^2_{in}\Big)H_{\bar g} + 2h_{nn}\Big(\nabla_i h_n{}^i - \frac{1}{2}\nabla_n h_i{}^i\Big)\Big)f\Big]d\sigma_{\bar g}$$
$$-\int_\Sigma\Big[\Big(2h^2_{nn} + \sum_{i=1}^{n-1}h^2_{in}\Big)\partial_n f + 2h_{nn}\sum_{i=1}^{n-1}h_{in}\partial_i f\Big]d\sigma_{\bar g}.$$

In general, geometric functionals are invariant under actions of diffeomorphisms and it would cause degenerations on their second variations. In order to get rid of these degenerations, we need to find a metric modulo diffeomorphisms. This is usually referred to be *gauge fixing* and it can be obtained by applying basic elliptic theory and the implicit function theorem. For manifolds with boundary, this can be achieved if one poses appropriate boundary conditions.

Lemma 4.3 [Brendle and Marques 2011, Proposition 11]. *Suppose $(\Omega^n, \bar g)$ is a compact Riemannian manifold with boundary. Fix a real number $p > n$. Then there exists a constant $\varepsilon_1 > 0$ such that for a metric g on Ω with*

$$g|_{T\partial\Omega} = \bar g|_{T\partial\Omega}$$

and

$$\|g - \bar g\|_{W^{2,p}(\Omega,\bar g)} < \varepsilon_1,$$

there exists a diffeomorphism $\varphi : \Omega \to \Omega$ such that $\varphi|_{\partial\Omega} = \mathrm{id}$ and $h := \varphi^*g - \bar g \in \ker\delta_{\bar g}$. Moreover,

$$\|h\|_{W^{2,p}(\Omega,\bar g)} \le N\|g - \bar g\|_{W^{2,p}(\Omega,\bar g)}$$

for some constant $N > 0$ that depends only on $(\Omega, \bar g)$.

In particular, we take Ω to be a geodesic ball $B_r(p)$ at an interior point $p \in M$ with radius $r > 0$.

Proposition 4.4. *Suppose $(M^n, \bar g, \kappa, f)$ is a V-static space and $p \in M$ is an interior point. Then there is a constant $\varepsilon_1 > 0$ such that for any metric g on $B_r(p)$ satisfying*

- *$R_g \ge R_{\bar g}$ in $B_r(p)$,*
- *$H_g \ge H_{\bar g}$ on $\partial B_r(p)$,*
- *$g|_{T\partial B_r(p)} = \bar g|_{T\partial B_r(p)}$,*
- *$\|g - \bar g\|_{C^2(B_r(p),\bar g)} < \varepsilon_1$,*

we can find a diffeomorphism $\varphi \in \mathscr{D}(B_r(p))$ such that $\varphi|_{\partial B_r(p)} = \mathrm{id}$ and

$$h := \varphi^*g - \bar g \in \ker\delta_{\bar g}$$

satisfies $|h|_{\bar g} < \frac{1}{2}$ in $B_r(p)$, $h|_{T\partial B_r(p)} \equiv 0$ on $\partial B_r(p)$ and

$$\|h\|_{C^2(B_r(p),\bar g)} \le N\|g - \bar g\|_{C^2(B_r(p),\bar g)}$$

for some constant $N > 0$ depending only on $(B_r(p), \bar{g})$. Additionally, we have

- $R_{\varphi^* g} \geq R_{\bar{g}}$ *in* $B_r(p)$,
- $H_{\varphi^* g} \geq H_{\bar{g}}$ *on* $\partial B_r(p)$.

Proof. The existence of a constant ε_1 and diffeomorphism φ is a straightforward application of Lemma 4.3. Furthermore, we have

- $R_{\varphi^* g} = R_g \circ \varphi \geq R_{\bar{g}}$ in $B_r(p)$,
- $H_{\varphi^* g} = H_g \circ \varphi = H_g \geq H_{\bar{g}}$ on $\partial B_r(p)$,

because of the fact that the scalar curvature $R_{\bar{g}}$ is a constant on M (see Remark 1.1) and $\varphi|_{\partial B_r(p)} = \mathrm{id}$. □

Let $\hat{g}_h = \bar{g} + h$ be a metric on $B_r(p)$, where $h \in S_2(B_r(p))$ satisfies $|h|_{\bar{g}} < \frac{1}{2}$ and $h|_{T\partial B_r(p)} \equiv 0$. From Propositions 4.1 and 4.2, the remainder of the expansion for $\mathscr{F}_{\Omega,\bar{g}}$ up to second order can be written as

$$\begin{aligned} r_{B_r(p),\bar{g}}[h] &:= \mathscr{F}_{B_r(p),\bar{g}}[\hat{g}_h] - \mathscr{F}_{B_r(p),\bar{g}}[\bar{g}] - D\mathscr{F}_{B_r(p),\bar{g}} \cdot h - \tfrac{1}{2} D^2 \mathscr{F}_{B_r(p),\bar{g}} \cdot (h,h) \\ &= \int_{B_r(p)} (R_{\hat{g}_h} - R_{\bar{g}}) f \, dv_{\bar{g}} - 2\kappa (V_{B_r(p)}(\hat{g}_h) - V_{B_r(p)}(\bar{g})) + I_{B_r(p)}[h] + I_{\partial B_r(p)}[h], \end{aligned} \tag{4-3}$$

where

$$I_{B_r(p)}[h] := \frac{1}{4} \int_{B_r(p)} \Big[(|\nabla_{\bar{g}} h|^2_{\bar{g}} + |d(\mathrm{tr}_{\bar{g}}\, h)|^2 - 2\mathscr{R}_{\bar{g}}(h,h)) f + \frac{n+3}{n-1} (\mathrm{tr}_{\bar{g}}\, h)^2 \kappa \Big] \, dv_{\bar{g}}$$

and

$$\begin{aligned} I_{\partial B_r(p)}[h] := &\int_{\partial B_r(p)} \Big[2(H_{\hat{g}_h} - H_{\bar{g}}) + \frac{1}{2} A^{ij}_{\bar{g}} h_{in} h_{jn} - \frac{1}{4} \Big(h^2_{nn} - 2 \sum_{i=1}^{n-1} h^2_{in} \Big) H_{\bar{g}} + h_{nn} \Big(\nabla_i h_n{}^i - \frac{1}{2} \nabla_n h_i{}^i \Big) \Big] f \, d\sigma_{\bar{g}} \\ &+ \int_{\partial B_r(p)} \Big[\Big(h^2_{nn} + \frac{1}{2} \sum_{i=1}^{n-1} h^2_{in} \Big) \partial_n f + h_{nn} \sum_{i=1}^{n-1} h_{in} \partial_i f \Big] \, d\sigma_{\bar{g}}. \end{aligned}$$

The estimate for the remainder $r_{B_r(p),\bar{g}}[h]$ plays a key role in our proof. It mainly relies on estimates for lower bounds of integrals $I_{B_r(p)}$ and $I_{\partial B_r(p)}$.

The estimate for a lower bound of interior integral $I_{B_r(p)}$ is essentially due to the solution of the variational problem

$$\mu(\Omega, \bar{g}) = \inf \left\{ \frac{\int_\Omega |\nabla_{\bar{g}} h|^2_{\bar{g}} \, dv_{\bar{g}}}{\int_\Omega |h|^2_{\bar{g}} \, dv_{\bar{g}}} : h \in S_2(\Omega),\ h \not\equiv 0 \text{ and } h|_{T\partial\Omega} \equiv 0 \right\}.$$

A basic estimate was obtained by Qing and the author in [Qing and Yuan 2016, Lemma 3.7]:

Lemma 4.5. *Suppose $(M^n, \bar{g})$ is a Riemannian manifold with dimension $n \geq 3$ and $B_r(p)$ is a geodesic ball of radius r centered at any interior point $p \in M$. Then there are positive constants $\bar{r}$ and c_0 such that*

$$\mu(B_r(p), \bar{g}) \geq \frac{c_0}{r^2} \tag{4-4}$$

for all $0 < r < \bar{r}$.

From this, we are ready to obtain an estimate for a lower bound of $I_{B_r(p)}$.

Proposition 4.6. *Suppose $p \in M$ is an interior point with $f(p) > 0$. Then there is a constant $r_1 > 0$ such that*

$$f(x) > 0$$

for all $x \in \overline{B_{r_1}(p)} \subseteq M$. Furthermore, for all $r \in (0, r_1)$ and any $h \in S_2(B_r(p))$ with $h|_{T\partial B_r(p)} \equiv 0$,

$$I_{B_r(p)}[h] \geq \frac{1}{8}\big(\inf_{B_r(p)} f\big)\|h\|^2_{W^{1,2}(B_r(p),\bar{g})}. \tag{4-5}$$

Proof. By continuity, we can choose a constant $r_1' > 0$ such that $f(x) > 0$ for all $x \in \overline{B_{r_1'}(p)}$.

It is straightforward to see that

$$|\mathscr{R}_{\bar{g}}(h,h)| = \left|\langle \mathrm{Rm}_{\bar{g}} \cdot h, h\rangle_{\bar{g}} + 2(\mathrm{Ric}_{\bar{g}} \cdot h)(\mathrm{tr}_{\bar{g}}\, h) - \frac{2R_{\bar{g}}}{n-1}(\mathrm{tr}_{\bar{g}}\, h)^2\right| \leq \Lambda_{r_1'}|h|^2_{\bar{g}}$$

on $B_{r_1'}(p)$, where $\Lambda_{r_1'} = \Lambda(n, \bar{g}, \|\mathrm{Rm}_{\bar{g}}\|_{C^0(B_{r_1'}(p),\bar{g})})$ is a positive constant independent of h. Thus for any $r < r_1'$ and $h \in S_2(B_r(p))$ with $h|_{T\partial B_r(p)} \equiv 0$, we have

$$\begin{aligned}
I_{B_r(p)}[h] &\geq \frac{1}{4}\int_{B_r(p)} [(|\nabla_{\bar{g}} h|^2_{\bar{g}} - 2|\mathscr{R}_{\bar{g}}(h,h)|) f - 3n|\kappa|\,|h|^2_{\bar{g}}]\, dv_{\bar{g}} \\
&\geq \frac{1}{4}\int_{B_r(p)} \big[\big(\inf_{B_r(p)} f\big)|\nabla_{\bar{g}} h|^2_{\bar{g}} - \big(2\Lambda_{r_1'}\big(\sup_{B_r(p)} f\big) + 3n|\kappa|\big)|h|^2_{\bar{g}}\big]\, dv_{\bar{g}} \\
&= \frac{1}{8}\big(\inf_{B_r(p)} f\big)\|h\|^2_{W^{1,2}(B_r(p),\bar{g})} + \frac{1}{8}\big(\inf_{B_r(p)} f\big)\int_{B_r(p)} [|\nabla_{\bar{g}} h|^2_{\bar{g}} - \mu_r|h|^2_{\bar{g}}]\, dv_{\bar{g}},
\end{aligned}$$

where

$$\mu_r := \frac{4\Lambda_{r_1'}(\sup_{B_r(p)} f) + (\inf_{B_r(p)} f) + 6n|\kappa|}{\inf_{B_r(p)} f} \leq \frac{(4\Lambda_{r_1'} + 1)(\sup_{B_{r_1'}(p)} f) + 6n|\kappa|}{\inf_{B_{r_1'}(p)} f} := \bar{\mu}_{r_1'}.$$

Applying Lemma 4.5, we can choose a positive constant $r_1 < r_1'$ sufficiently small such that

$$\int_{B_r(p)} |\nabla_{\bar{g}} h|^2_{\bar{g}}\, dv_{\bar{g}} \geq \bar{\mu}_{r_1'} \int_{B_r(p)} |h|^2_{\bar{g}}\, dv_{\bar{g}}$$

for all $r \in (0, r_1)$. Therefore

$$I_{B_r(p)}[h] \geq \frac{1}{8}\big(\inf_{B_r(p)} f\big)\|h\|^2_{W^{1,2}(B_r(p),\bar{g})}$$

for any $r \in (0, r_1)$. $\square$

For a lower bound estimate for the boundary integral $I_{\partial B_r(p)}$ we have the following.

Proposition 4.7. *Suppose $p \in M$ is an interior point with $f(p) > 0$. Then there is a constant $r_2 > 0$ such that*

$$f(x) > 0$$

for all $x \in \overline{B_{r_2}(p)} \subseteq M$. Furthermore, for all $r \in (0, r_2)$ and any metric $\hat{g}_h := \bar{g} + h$ in $B_r(p)$ satisfying

- *$h \in S_2(B_r(p))$ with $|h|_{\bar{g}} < \frac{1}{2}$ and $h|_{T\partial B_r(p)} \equiv 0$,*
- *$H_{\hat{g}_h} \geq H_{\bar{g}}$ on $\partial B_r(p)$,*

we have

$$I_{\partial B_r(p)}[h] \geq -C_0\big(\sup_{B_r(p)} f\big)\|h\|_{C^1(B_r(p),\bar{g})}\|h\|^2_{W^{1,2}(B_r(p),\bar{g})}, \tag{4-6}$$

where $C_0 > 0$ is a constant depending only on $(B_r(p), \bar{g})$.

Proof. By continuity, we can choose a constant $r_2' > 0$ such that $f(x) > 0$ for all $x \in \overline{B_{r_2'}(p)}$.

As observed in [Brendle and Marques 2011], for all $r \in (0, r_2')$ and any metric $\hat{g}_h = \bar{g} + h$ satisfying $h \in S_2(B_r(p))$ with $|h|_{\bar{g}} < \frac{1}{2}$ and $h|_{T\partial B_r(p)} \equiv 0$, we have

$$h_{nn}(H_{\hat{g}_h} - H_{\bar{g}}) = \tfrac{1}{2}h_{nn}^2 H_{\bar{g}} - h_{nn}\big(\nabla_i h_n{}^i - \tfrac{1}{2}\nabla_n h_i{}^i\big) + F_{\bar{g}}(h)$$

due to Lemma 3.2, where the tail term $F_{\bar{g}}(h)$ satisfies

$$|F_{\bar{g}}(h)|_{\bar{g}} \leq \tilde{C}_1 |h|^2_{\bar{g}}(|\nabla_{\bar{g}} h|_{\bar{g}} + |A_{\bar{g}}|_{\bar{g}}|h|_{\bar{g}}),$$

and $\tilde{C}_1 > 0$ is a constant depending only on the dimension n. From this,

$$\begin{aligned} I_{\partial B_r(p)}[h] = \int_{\partial B_r(p)} \bigg[(2 - h_{nn})(H_{\hat{g}} - H_{\bar{g}}) + \frac{1}{2}A_{\bar{g}}^{ij} h_{in}h_{jn} + \frac{1}{4}\bigg(h_{nn}^2 + 2\sum_{i=1}^{n-1} h_{in}^2\bigg) H_{\bar{g}}\bigg] f\, d\sigma_{\bar{g}} \\ + \int_{\partial B_r(p)} \bigg[\bigg(h_{nn}^2 + \frac{1}{2}\sum_{i=1}^{n-1} h_{in}^2\bigg)\partial_n f + h_{nn}\sum_{i=1}^{n-1} h_{in}\partial_i f\bigg] d\sigma_{\bar{g}} + \tilde{F}_{\bar{g}}(h), \end{aligned}$$

where the tail term $\tilde{F}_{\bar{g}}(h)$ satisfies

$$|\tilde{F}_{\bar{g}}(h)| \leq \tilde{C}_2\big(\sup_{B_r(p)} f\big) \int_{\partial B_r(p)} |h|^2_{\bar{g}}(|\nabla_{\bar{g}} h|_{\bar{g}} + |A_{\bar{g}}|_{\bar{g}}|h|_{\bar{g}})\, dv_{\bar{g}}$$

for a constant $\tilde{C}_2 > 0$ depending only on the dimension n.

For $r > 0$ sufficiently small, it is well known that the second fundamental form and mean curvature of the geodesic sphere $\partial B_r(p)$ behave similarly to round spheres in Euclidean space (see Exercise 1.123 in [Chow et al. 2006]):

$$A_{ij}^{\bar{g}} = \frac{1}{r}\bar{g}_{ij} + O(r) \qquad \text{and} \qquad H_{\bar{g}} = \frac{n-1}{r} + O(r)$$

on $\partial B_r(p)$. Thus we can choose $r_2'' \in (0, r_2')$ such that

$$A_{ij}^{\bar{g}} \geq \frac{1}{2r}\bar{g}_{ij} \qquad \text{and} \qquad H_{\bar{g}} \geq \frac{n-1}{2r}$$

for any geodesic sphere $\partial B_r(p)$ with $r < r_2''$.

For $r \in (0, r_2'')$, we have

$$\begin{aligned} I_{\partial B_r(p)}[h] &\geq \frac{1}{2}\int_{\partial B_r(p)} \bigg[\frac{1}{4r}\bigg((n-1)h_{nn}^2 + 2n\sum_{i=1}^{n-1} h_{in}^2\bigg) f - \bigg(3h_{nn}^2 + n\sum_{i=1}^{n-1} h_{in}^2\bigg)|\nabla_{\bar{g}} f|_{\bar{g}}\bigg] d\sigma_{\bar{g}} + \tilde{F}_{\bar{g}}(h) \\ &= \frac{1}{2}\int_{\partial B_r(p)} \bigg[3\bigg(\frac{n-1}{12r} - \frac{|\nabla_{\bar{g}} f|_{\bar{g}}}{f}\bigg)h_{nn}^2 + n\bigg(\frac{1}{2r} - \frac{|\nabla_{\bar{g}} f|_{\bar{g}}}{f}\bigg)\sum_{i=1}^{n-1} h_{in}^2\bigg] f\, d\sigma_{\bar{g}} + \tilde{F}_{\bar{g}}(h). \end{aligned}$$

Since f is positively lower bounded and $|\nabla_{\bar{g}} f|_{\bar{g}}$ is upper bonded on $B_{r_2''}(p)$, we can pick a constant $r_2 \in (0, r_2'')$ such that

$$\frac{|\nabla_{\bar{g}} f|_{\bar{g}}}{f} \leq \min\left\{\frac{n-1}{12r}, \frac{1}{2r}\right\}$$

holds in $B_r(p)$ for any $r \in (0, r_2)$ and hence

$$I_{\partial B_r(p)} \geq \tilde{F}_{\bar{g}}(h) \geq -\tilde{C}_3\big(\sup_{B_r(p)} f\big)\|h\|_{C^1(\partial B_r(p),\bar{g})}\|h\|^2_{L^2(\partial B_r(p),\bar{g})}$$

for any $r \in (0, r_2)$, where $\tilde{C}_3 > 0$ is a constant depending only on n and r.

Recall the *Sobolev trace inequality*

$$\|h\|^2_{L^2(\partial B_r(p),\bar{g})} \leq \theta_0 \ \|h\|^2_{W^{1,2}(B_r(p),\bar{g})},$$

where $\theta_0 > 0$ is a constant depending only on $(B_r(p), \bar{g})$. Therefore the estimate

$$I_{\partial B_r(p)} \geq -C_0\big(\sup_{B_r(p)} f\big)\|h\|_{C^1(B_r(p),\bar{g})}\|h\|^2_{W^{1,2}(B_r(p),\bar{g})}$$

holds for any $r \in (0, r_2)$, where $C_0 := \theta_0\tilde{C}_3 > 0$ is a constant depending only on $(B_r(p), \bar{g})$. □

Now we are ready to prove the main theorem in this section.

Proof of Theorem A. Let

$$r_0 := \min\{r_1, r_2\} > 0,$$

where r_1 and r_2 are given by Propositions 4.6 and 4.7.

For all $r \in (0, r_0)$, applying Proposition 4.4, we can find a constant $\varepsilon_1 > 0$ such that for any metric g on $B_r(p) \subset M$ satisfying

- $R_g \geq R_{\bar{g}}$ in $B_r(p)$,
- $H_g \geq H_{\bar{g}}$ on $\partial B_r(p)$,
- $g|_{T\partial B_r(p)} = \bar{g}|_{T\partial B_r(p)}$,
- $\|g - \bar{g}\|_{C^2(B_r(p),\bar{g})} < \varepsilon_1$,

there is a diffeomorphism $\varphi \in \mathscr{D}(B_r(p))$ such that $\varphi|_{\partial B_r(p)} = \mathrm{id}$ and

$$h := \varphi^* g - \bar{g} \in \ker \delta_{\bar{g}}$$

satisfies $|h|_{\bar{g}} < \frac{1}{2}$ in $B_r(p)$, $h|_{T\partial B_r(p)} \equiv 0$ on $\partial B_r(p)$ and

$$\|h\|_{C^2(B_r(p),\bar{g})} \leq N\|g - \bar{g}\|_{C^2(B_r(p),\bar{g})}$$

for some constant $N > 0$ depending only on $(B_r(p), \bar{g})$. Additionally, we have

- $R_{\varphi^* g} \geq R_{\bar{g}}$ in $B_r(p)$,
- $H_{\varphi^* g} \geq H_{\bar{g}}$ on $\partial B_r(p)$.

Fix an $r \in (0, r_0)$ and assume the contrary of the claimed volume comparison:

$$\kappa(V_{B_r(p)}(g) - V_{B_r(p)}(\bar{g})) \leq 0, \tag{4-7}$$

which implies

$$\kappa(V_{B_r(p)}(\varphi^* g) - V_{B_r(p)}(\bar{g})) \leq 0.$$

By Propositions 4.6 and 4.7, the lower bound estimate for the remainder is

$$\begin{aligned} r_{B_r(p),\bar{g}}[h] &= \mathscr{F}_{B_r(p),\bar{g}}[\varphi^* g] - \mathscr{F}_{B_r(p),\bar{g}}[\bar{g}] - D\mathscr{F}_{B_r(p),\bar{g}} \cdot h - \tfrac{1}{2} D^2 \mathscr{F}_{B_r(p),\bar{g}} \cdot (h, h) \\ &= \int_{B_r(p)} (R_{\varphi^* g} - R_{\bar{g}}) f \, dv_{\bar{g}} - 2\kappa(V_{B_r(p)}(\varphi^* g) - V_{B_r(p)}(\bar{g})) + I_{B_r(p)}[h] + I_{\partial B_r(p)}[h] \\ &\geq \Big(\tfrac{1}{8}\big(\inf_{B_r(p)} f\big) - C_0\big(\sup_{B_r(p)} f\big)\|h\|_{C^1(B_r(p),\bar{g})}\Big)\|h\|^2_{W^{1,2}(B_r(p),\bar{g})}. \end{aligned}$$

On the other hand, if we write

$$\tau_r := \max\Big\{\sup_{B_r(p)} f,\ \sup_{B_r(p)} |\nabla_{\bar{g}} f|_{\bar{g}}\Big\},$$

then the upper bound of the remainder can be estimated using Taylor's formula:

$$\begin{aligned} r_{B_r(p),\bar{g}}[h] &= \tfrac{1}{6} D^3 \mathscr{F}_{B_r(p),\bar{g}+\xi h} \cdot (h, h, h) \\ &\leq C_1 \tau_r \int_{B_r(p)} |h|_{\bar{g}} (|\nabla_{\bar{g}} h|^2_{\bar{g}} + |h|^2_{\bar{g}}) \, dv_{\bar{g}} + C_2 \tau_r \int_{\partial B_r(p)} |h|^2_{\bar{g}} (|\nabla_{\bar{g}} h|_{\bar{g}} + |A_{\bar{g}}|_{\bar{g}} |h|_{\bar{g}}) \, dv_{\bar{g}} \\ &\leq C_1 \tau_r \|h\|_{C^0(B_r(p),\bar{g})} \|h\|^2_{W^{1,2}(B_r(p),\bar{g})} + C_3 \tau_r \|h\|_{C^1(B_r(p),\bar{g})} \|h\|^2_{L^2(\partial B_r(p),\bar{g})}, \end{aligned}$$

where $\xi \in (0, 1)$ is a constant and C_1, C_2, C_3 are positive constants depending only on $(B_r(p), \bar{g})$. Recall again the *Sobolev trace inequality*

$$\|h\|^2_{L^2(\partial B_r(p),\bar{g})} \leq \theta_0 \ \|h\|^2_{W^{1,2}(B_r(p),\bar{g})},$$

where $\theta_0 > 0$ is a constant depending only on $(B_r(p), \bar{g})$. From this we obtain

$$r_{B_r(p),\bar{g}}[h] \leq C_0' \tau_r \|h\|_{C^1(B_r(p),\bar{g})} \|h\|^2_{W^{1,2}(B_r(p),\bar{g})},$$

where $C_0' = C_1 + \theta_0 C_3$ is a positive constant depending only on $(B_r(p), \bar{g})$.

Combining both lower and upper bound estimates of $r_{B_r(p),\bar{g}}$, we obtain

$$\Big(\tfrac{1}{8}\big(\inf_{B_r(p)} f\big) - \Big(C_0\big(\sup_{B_r(p)} f\big) + C_0'\tau_r\Big)\|h\|_{C^1(B_r(p),\bar{g})}\Big)\|h\|^2_{W^{1,2}(B_r(p),\bar{g})} \leq 0. \tag{4-8}$$

Take

$$\varepsilon_0 := \frac{1}{N} \min\Big\{\varepsilon_1,\ \tfrac{1}{8}\Big(C_0\big(\sup_{B_r(p)} f\big) + C_0'\tau_r\Big)^{-1}\big(\inf_{B_r(p)} f\big)\Big\}.$$

Then for the metric g satisfying

$$\|g - \bar{g}\|_{C^2(B_r(p),\bar{g})} < \varepsilon_0$$

we have

$$\|h\|_{C^1(B_r(p),\bar{g})} \leq N\|g - \bar{g}\|_{C^2(B_r(p),\bar{g})} < N\varepsilon_0 < \tfrac{1}{8}\Big(C_0\big(\sup_{B_r(p)} f\big) + C_0'\tau_r\Big)^{-1}\big(\inf_{B_r(p)} f\big).$$

According to inequality (4-8), we see h vanishes identically on $B_r(p)$ and hence $\varphi^*g = \bar{g}$, which shows that $\varphi : B_r(p) \to B_r(p)$ has to be an isometry. Therefore the reverse of inequality (4-7) holds:

$$\kappa(V_{B_r(p)}(g) - V_{B_r(p)}(\bar{g})) \geq 0. \tag{4-9}$$

That is, the following volume comparison holds:

- if $\kappa < 0$, then
$$V_{B_r(p)}(g) \leq V_{B_r(p)}(\bar{g});$$
- if $\kappa > 0$, then
$$V_{B_r(p)}(g) \geq V_{B_r(p)}(\bar{g});$$

with equality holding in either case if and only if the metric g is isometric to $\bar{g}$. □

5. Volume comparison for closed Einstein manifolds

Suppose $(M^n, \bar{g}, f, \kappa)$ is a closed V-static manifold. Then the functional $\mathscr{F}_{M,\bar{g}}$ introduced in the previous section can be simplified as

$$\mathscr{F}_{M,\bar{g}}[g] = \int_M R(g) f \, dv_{\bar{g}} - 2\kappa V_M(g). \tag{5-1}$$

According to Proposition 4.1, the metric $\bar{g}$ is still a critical point of $\mathscr{F}_{M,\bar{g}}$. However, it is obvious that this functional is not compatible with actions of dilations, which would cause subtle issues in its second variation. Geometrically speaking, dilations introduce additional degeneracy besides actions of diffeomorphisms, since they make no essential change to the geometry of the manifold. In order to obtain volume comparison for closed manifolds, we need to construct a new functional instead, which is invariant under dilations.

Definition 5.1. Suppose $(M^n, \bar{g}, f, \kappa)$ is an n-dimensional closed V-static manifold. We define the functional

$$\mathscr{G}_{M,\bar{g}}[g] := (V_M(g))^{2/n} \int_M R(g) f \, dv_{\bar{g}} \tag{5-2}$$

for any Riemannian metric g on M.

Obviously, this functional is dilation-invariant:

$$\mathscr{G}_{M,\bar{g}}[c^2 g] = (V_M(c^2 g))^{2/n} \int_M R(c^2 g) f \, dv_{\bar{g}} = \mathscr{G}_{M,\bar{g}}[g]$$

for any constant $c \neq 0$.

Now we focus on a special type of V-static metrics: Einstein metrics. According to the V-static equation (1-1), we get

$$\gamma_{\bar{g}}^* 1 = -\operatorname{Ric}_{\bar{g}} = \kappa \bar{g}$$

by taking the function f to be constantly 1 on M. This means $(M^n, \bar{g}, 1, \kappa)$ is a V-static space if and only if the metric $\bar{g}$ is an Einstein metric with scalar curvature $R_{\bar{g}} = -n\kappa$. Moreover, if we write

$$\lambda := \frac{R_{\bar{g}}}{n(n-1)}, \tag{5-3}$$

then the Ricci curvature tensor is given by

$$\mathrm{Ric}_{\bar g} = (n-1)\lambda \bar g$$

and

$$\kappa = -(n-1)\lambda.$$

As a functional designed for V-static metrics, $\mathscr{G}_{M,\bar g}$ shares analogous variational properties with $\mathscr{F}_{M,\bar g}$.

Proposition 5.2. *Suppose $(M^n, \bar g)$ is a closed Einstein manifold with Ricci curvature tensor*

$$\mathrm{Ric}_{\bar g} = (n-1)\lambda \bar g.$$

Then the metric $\bar g$ is a critical point of the functional $\mathscr{G}_{M,\bar g}$.

Proof. From Proposition 3.4 and Lemma 3.3,

$$\begin{aligned} D\mathscr{G}_{M,\bar g}\cdot h &= (V_M(\bar g))^{2/n}\int_M (DR_{\bar g}\cdot h)\,dv_{\bar g} + \frac{2}{n}(V_M(\bar g))^{(2/n)-1}(DV_{M,\bar g}\cdot h)\int_M R_{\bar g}\,dv_{\bar g}\\ &= (V_M(\bar g))^{2/n}\left[\int_M (\gamma_{\bar g}^* 1)\ \,dv_{\bar g} + \frac{1}{n}R_{\bar g}\int_M (\mathrm{tr}_{\bar g}\,h)\,dv_{\bar g}\right]\\ &= -(V_M(\bar g))^{2/n}\int_M \langle \mathrm{Ric}_{\bar g} - (n-1)\lambda \bar g, h\rangle_{\bar g}\,dv_{\bar g} = 0, \end{aligned}$$

for any $h \in S_2(M)$. □

For the second variation, we have the following.

Proposition 5.3. *Suppose (M^n, g) is an Einstein manifold with Ricci curvature tensor*

$$\mathrm{Ric}_{\bar g} = (n-1)\lambda \bar g.$$

Then

$$D^2\mathscr{G}_{M,\bar g}\cdot(h,h) = -\frac{1}{2}(V_M(\bar g))^{2/n}\int_M\Big[-\langle h_{\mathrm{TT}}, \Delta_E^{\bar g} h_{\mathrm{TT}}\rangle_{\bar g} + \frac{(n-1)(n+2)}{n^2}(|d(\mathrm{tr}_{\bar g}\,h)|_{\bar g}^2 - n\lambda(\mathrm{tr}_{\bar g}\,h - \overline{\mathrm{tr}_{\bar g}\,h})^2)\Big]\,dv_{\bar g}$$

for any $h = h_{\mathrm{TT}} + \frac{1}{n}(\mathrm{tr}_{\bar g}\,h)\bar g \in S_{2,\bar g}^{\mathrm{TT}} \oplus (C^\infty(M)\cdot \bar g)$.

Proof. From Lemmas 3.1 and 3.3 and Corollary 3.7 we obtain

$$\begin{aligned} &D^2\mathscr{G}_{M,\bar g}\cdot(h,h)\\ &\quad= \frac{2}{n}(V_M(\bar g))^{(2/n)-1}(D^2V_{M,\bar g}\cdot(h,h))\int_M R_{\bar g}\,dv_{\bar g} + \frac{4}{n}(V_M(\bar g))^{(2/n)-1}(DV_{M,\bar g}\cdot h)\int_M (DR_{\bar g}\cdot h)\,dv_{\bar g}\\ &\qquad - \frac{2(n-2)}{n^2}(V_M(\bar g))^{(2/n)-2}(DV_{M,\bar g}\cdot h)^2\int_M R_{\bar g}\,dv_{\bar g} + (V_M(\bar g))^{2/n}\int_M (D^2R_{\bar g}\cdot(h,h))\,dv_{\bar g}\\ &\quad= -\frac{1}{2}(V_M(\bar g))^{2/n}\int_M (-\langle h, \Delta_E^{\bar g} h\rangle_{\bar g} + \frac{n^2-2}{n^2}|d(\mathrm{tr}_{\bar g}\,h)|_{\bar g}^2 - (n-1)\lambda(\mathrm{tr}_{\bar g}\,h)^2)\,dv_{\bar g}\\ &\qquad - \frac{(n-1)(n+2)}{2n}\lambda(V_M(\bar g))^{2/n}\int_M (\overline{\mathrm{tr}_{\bar g}\,h})^2\,dv_{\bar g}. \end{aligned}$$

Now the decomposition

$$h = h_{\mathrm{TT}} + \frac{1}{n}(\mathrm{tr}_{\bar g}\, h)\bar g$$

implies

$$(D^2\mathscr{G}_{M,\bar g})\cdot(h,h)$$
$$= -\frac{1}{2}(V_M(\bar g))^{2/n}\int_M\Big[-\langle h_{\mathrm{TT}}, \Delta_E^{\bar g} h_{\mathrm{TT}}\rangle_{\bar g} + \frac{(n-1)(n+2)}{n^2}\big(|d(\mathrm{tr}_{\bar g}\, h)|_{\bar g}^2 - n\lambda(\mathrm{tr}_{\bar g}\, h - \overline{\mathrm{tr}_{\bar g}\, h})^2\big)\Big]\, dv_{\bar g}. \qquad \square$$

As a key step of the proof for our volume comparison theorem, we need to give a characterization of the second variation of the functional $\mathscr{G}_{M,\bar g}$ at $\bar g$. This is closely related to spectrum problems of two operators: one is about the Einstein operator and can be characterized by the stability of Einstein metrics, the other is about the Laplace–Beltrami operator whose eigenvalue estimate is given by the well-known *Lichnerowicz–Obata theorem*; see Theorem 5.1 in [Li 2012].

Lemma 5.4 (Lichnerowicz–Obata's eigenvalue estimate). *Suppose $(M^n, \bar g)$ is an n-dimensional closed Riemannian manifold with Ricci curvature tensor*

$$\mathrm{Ric}_{\bar g} \geq (n-1)\lambda \bar g,$$

where $\lambda > 0$ is a constant. Then for any function $u \in C^\infty(M)$ that is not identically a constant, we have

$$\int_M |du|^2\, dv_{\bar g} \geq n\lambda \int_M (u-\bar u)^2\, dv_{\bar g}, \tag{5-4}$$

where equality holds if and only if $(M^n, \bar g)$ is isometric to the round sphere $\mathbb{S}^n(r)$ with radius $r = 1/\sqrt{\lambda}$ and u is a first eigenfunction of the Laplace–Beltrami operator.

Applying this to Proposition 5.3, immediately we get the nonpositive definite property of the second variation of $\mathscr{G}_{M,\bar g}$ at $\bar g$.

Proposition 5.5. *Suppose $(M^n, \bar g)$ is a closed stable Einstein manifold with Ricci curvature tensor*

$$\mathrm{Ric}_{\bar g} = (n-1)\lambda \bar g.$$

Then

$$D^2\mathscr{G}_{M,\bar g}\cdot(h,h) \leq 0$$

for any $h \in S_{2,\bar g}^{\mathrm{TT}}(M) \oplus (C^\infty(M)\cdot \bar g)$. Moreover, equality holds if and only if

- *$h \in \mathbb{R}\bar g \oplus \ker \Delta_E^{\bar g}$, when $(M, \bar g)$ is not isometric to the round sphere up to a rescaling of the metric,*
- *$h \in (\mathbb{R} \oplus E_{n\lambda})\bar g$, when $(M, \bar g)$ is isometric to the round sphere $\mathbb{S}^n(r)$ with radius $r = 1/\sqrt{\lambda}$,*

where

$$E_{n\lambda} := \{u \in C^\infty(\mathbb{S}^n(r)) : \Delta_{\mathbb{S}^n(r)} u + n\lambda u = 0\}$$

is the space of first eigenfunctions for the spherical metric.

Proof. Recall that the Einstein metric $\bar g$ is stable if and only if $(-\Delta_E^{\bar g})$ is a nonnegative operator. Then the conclusion follows by applying this fact and Lemma 5.4 to Proposition 5.3. $\square$

Intuitively speaking, a *slice* is a subset of metrics in the space of all Riemannian metrics which is transverse to the orbit of diffeomorphism actions. The following refined version of the slice theorem reveals the local structure of Einstein metrics in the space of all metrics. To the best of the author's knowledge, it does not appear in the literature. We hope it can be useful in problems involving Einstein metrics. The proof is standard; please refer to [Brendle and Marques 2011; Viaclovsky 2016].

Theorem 5.6 (Ebin–Palais slice theorem). *Suppose $(M^n, \bar{g})$ is a closed n-dimensional Einstein manifold with Ricci curvature tensor*

$$\mathrm{Ric}_{\bar{g}} = (n-1)\lambda \bar{g},$$

where $\lambda \in \mathbb{R}$ is a constant. Let $\mathcal{M}$ be the space of all Riemannian metrics on M. There exists a local slice $\mathcal{S}_{\bar{g}}$ though $\bar{g}$ in $\mathcal{M}$. That is, for a fixed real number $p > n$, one can find a constant $\varepsilon_1 > 0$ such that, for any metric $g \in \mathcal{M}$ with $\|g - \bar{g}\|_{W^{2,p}(M,\bar{g})} < \varepsilon_1$, there is a diffeomorphism $\varphi \in \mathscr{D}(M)$ with $\varphi^* g \in \mathcal{S}_{\bar{g}}$. Moreover, for a smooth local slice $\mathcal{S}_{\bar{g}}$, we have the decomposition

$$S_2(M) = T_{\bar{g}}\mathcal{S}_{\bar{g}} \oplus (T_{\bar{g}}\mathcal{S}_{\bar{g}})^{\perp},$$

where the tangent space of $\mathcal{S}_{\bar{g}}$ at $\bar{g}$ and its L^2-orthogonal complement are given by

$$T_{\bar{g}}\mathcal{S}_{\bar{g}} = S^{\mathrm{TT}}_{2,\bar{g}}(M) \oplus (C^{\infty}(M) \cdot \bar{g})$$

and

$$(T_{\bar{g}}\mathcal{S}_{\bar{g}})^{\perp} = \{\mathcal{L}_{\bar{g}}(X) : \langle X, \nabla_{\bar{g}} u\rangle_{L^2(M,\bar{g})} = 0 \text{ for all } u \in C^{\infty}(M)\}$$

when $(M^n, \bar{g})$ is not isometric to the round sphere $\mathbb{S}^n(r)$ up to a scaling, and

$$T_{\bar{g}}\mathcal{S}_{\bar{g}} = S^{\mathrm{TT}}_{2,\bar{g}}(M) \oplus (E^{\perp}_{n\lambda} \cdot \bar{g})$$

and

$$(T_{\bar{g}}\mathcal{S}_{\bar{g}})^{\perp} = \{\mathcal{L}_{\bar{g}}(X) : \langle X, \nabla_{\bar{g}} u\rangle_{L^2(M,\bar{g})} = 0 \text{ for all } u \in E^{\perp}_{n\lambda}\}$$

when $(M^n, \bar{g})$ is isometric to the round sphere $\mathbb{S}^n(r)$ with $r = 1/\sqrt{\lambda}$. Here

$$E_{n\lambda} = \{u \in C^{\infty}(\mathbb{S}^n(r)) : \Delta_{\mathbb{S}^n(r)} u + n\lambda u = 0\}$$

is the space of first eigenfunctions for the spherical metric.

Now we restrict the functional $\mathscr{G}_{M,\bar{g}}$ on a local slice $\mathcal{S}_{\bar{g}}$ and denote it by

$$\mathscr{G}^{\mathcal{S}}_{M,\bar{g}} := \mathscr{G}_{M,\bar{g}}|_{\mathcal{S}}.$$

In order to investigate the local behavior of $\mathscr{G}^{\mathcal{S}}_{M,\bar{g}}$ near $\bar{g}$, we need the following *Morse lemma on Banach manifolds*.

Lemma 5.7 (Morse lemma [Fischer and Marsden 1975]). *Let $\mathcal{P}$ be a Banach manifold and $F : \mathcal{P} \to \mathbb{R}$ a C^2-function. Suppose $\mathcal{Q} \subset \mathcal{P}$ is a submanifold, $F = 0$ and $dF = 0$ on $\mathcal{Q}$ and that there is a smooth normal bundle neighborhood of $\mathcal{Q}$ such that if $\mathcal{E}_x$ is the normal complement to $T_x\mathcal{Q}$ in $T_x\mathcal{P}$ then $d^2F(x)$ is weakly negative definite on $\mathcal{E}_x$ (i.e., $d^2F(x)(v, v) \le 0$ with equality only if $v = 0$). Let $\langle\,\cdot\,,\cdot\,\rangle_x$ be a weak Riemannian structure with a smooth connection and assume that F has a smooth $\langle\,\cdot\,,\cdot\,\rangle_x$-gradient, $Y(x)$.*

Assume $DY(x)$ maps $\mathcal{E}_x$ to $\mathcal{E}_x$ and is an isomorphism for $x \in \mathcal{Q}$. Then there is a neighborhood U of $\mathcal{Q}$ such that $y \in U$ and $F(y) \geq 0$ implies $y \in \mathcal{Q}$.

Applying it to our case, we obtain the following local rigidity result.

Proposition 5.8. *Suppose $(M^n, \bar{g})$ is a strictly stable Einstein manifold and $\mathcal{S}_{\bar{g}}$ is a local slice through $\bar{g}$. Then there is a neighborhood $U_{\bar{g}}$ of $\bar{g}$ in $\mathcal{S}_{\bar{g}}$ such that for any metric $\hat{g}_s \in U_{\bar{g}}$ satisfying*

$$\mathscr{G}^{\mathcal{S}}_{M,\bar{g}}[\hat{g}_s] \geq \mathscr{G}^{\mathcal{S}}_{M,\bar{g}}[\bar{g}],$$

there is a constant $c > 0$ such that $\hat{g}_s = c^2\bar{g}$.

Proof. Let

$$\tilde{\mathcal{Q}}_{\bar{g}} := \{g_s \in \mathcal{S}_{\bar{g}} : g_s \text{ is Einstein}\}$$

be the subset of the local slice $\mathcal{S}_{\bar{g}}$ consisting of Einstein metrics near the reference metric $\bar{g}$. By [Koiso 1980, Corollary 3.4], strict stability implies that $\bar{g}$ is rigid. That is, we can find a neighborhood $\tilde{U}_{\bar{g}} \subseteq \mathcal{S}_{\bar{g}}$ of $\bar{g}$ such that

$$\mathcal{Q}_{\bar{g}} := \tilde{\mathcal{Q}}_{\bar{g}} \cap \tilde{U}_{\bar{g}} = \{g_s \in \tilde{U}_{\bar{g}} : g_s = c^2\bar{g},\ c > 0\}.$$

In particular, the tangent space of $\mathcal{Q}_{\bar{g}}$ at $\bar{g}$ is given by

$$T_{\bar{g}}\mathcal{Q}_{\bar{g}} = \mathbb{R}\bar{g}$$

and its L^2-orthogonal complement in $T_{\bar{g}}\mathcal{S}_{\bar{g}}$ can be expressed as

$$\mathcal{E}_{\bar{g}} := (T_{\bar{g}}\mathcal{Q}_{\bar{g}})^{\perp} = S^{\mathrm{TT}}_{2,\bar{g}}(M) \oplus (\Psi_{\bar{g}}(M) \cdot \bar{g})$$

due to Theorem 5.6, where

$$\Psi_{\bar{g}}(M) = \left\{ u \in E^{\perp}_{n\lambda} : \int_M u \ dv_{\bar{g}} = 0 \right\}$$

if $\bar{g}$ is spherical and

$$\Psi_{\bar{g}}(M) = \left\{ u \in C^{\infty}(M) : \int_M u \, dv_{\bar{g}} = 0 \right\}$$

otherwise.

Consider a weak Riemannian structure on the local slice $\mathcal{S}_{\bar{g}}$,

$$\langle\!\langle \cdot , \cdot \rangle\!\rangle_{g_s} : T_{g_s}\mathcal{S}_{\bar{g}} \times T_{g_s}\mathcal{S}_{\bar{g}} \ \to \ \mathbb{R} \quad \text{for all } \ g_s \in \mathcal{S}_{\bar{g}},$$

which is defined to be

$$\langle\!\langle h, k \rangle\!\rangle_{g_s} := \int_M [\langle \nabla_{g_s} h, \nabla_{g_s} k\rangle_{g_s} + \langle h, k\rangle_{g_s}] \, dv_{g_s} = \int_M \langle (-\Delta_{g_s} + 1)h, k\rangle_{g_s} \, dv_{g_s}$$

for any $h, k \in T_{g_s}\mathcal{S}_{\bar{g}}$. According to [Ebin 1970] it has a smooth connection. The $\langle\!\langle \cdot , \cdot \rangle\!\rangle_{g_s}$-gradient of $\mathscr{G}^{\mathcal{S}}_{M,\bar{g}}$ is given by

$$Y(g_s) = P_{g_s}(-\Delta_{g_s} + 1)^{-1}\Big[(V_M(g_s))^{2/n}\Big(\gamma^*_{g_s} f_{g_s} + \frac{1}{n} g_s (V_M(g_s))^{-(n+2)/n}\mathscr{G}_{M,\bar{g}}[g_s]\Big)\Big],$$

where P_{g_S} is the orthogonal projection on $T_{g_S}\mathcal{S}_{\bar{g}}$ and f_{g_S} is a smooth function on M with $dv_{\bar{g}} = f_{g_S}\,dv_{g_S}$. Obviously, $Y(g_S)$ is a smooth vector field on $\mathcal{S}_{\bar{g}}$. For simplicity, we write

$$Z(g_S) := (V_M(g_S))^{2/n}\Big(\gamma^*_{g_S} f_{g_S} + \frac{1}{n} g_S (V_M(g_S))^{-(n+2)/n}\, \mathscr{G}_{M,\bar{g}}[g_S]\Big).$$

It is straightforward to see that $Z(\bar{g}) = 0$ and the linearization of Z at $\bar{g}$ is given by

$$(DZ_{\bar{g}})\cdot h = \frac{1}{2}(V_M(\bar{g}))^{2/n}\Big(\Delta^{\bar{g}}_E h_{\mathrm{TT}} + \frac{(n-1)(n+2)}{n^2}\bar{g}(\Delta_{\bar{g}} + n\lambda)(\operatorname{tr}_{\bar{g}} h - \overline{\operatorname{tr}_{\bar{g}} h})\Big) = D^2\mathscr{G}_{M,\bar{g}}\cdot(h,\cdot)$$

for any $h = h_{\mathrm{TT}} + \frac{1}{n}(\operatorname{tr}_{\bar{g}} h)\bar{g} \in \mathcal{E}_{\bar{g}}$. Thus

$$(DY_{\bar{g}})\cdot h = P_{\bar{g}}(-\Delta_{\bar{g}} + 1)^{-1}(D^2\mathscr{G}_{M,\bar{g}}\cdot(h,\cdot))$$

and $DY_{\bar{g}}$ is an isomorphism on $\mathcal{E}_{\bar{g}}$ due to the fact that $D^2\mathscr{G}^{\mathcal{S}}_{M,\bar{g}}$ is strictly negative definite on $\mathcal{E}_{\bar{g}}$ from Proposition 5.5.

Since the functional $\mathscr{G}^{\mathcal{S}}_{M,\bar{g}}$ is dilation-invariant, applying Lemma 5.7, we can find a neighborhood $U_{\bar{g}} \subseteq \mathcal{S}_{\bar{g}}$ of $\bar{g}$ such that for any $\hat{g}_S \in U_{\bar{g}}$ satisfying

$$\mathscr{G}^{\mathcal{S}}_{M,\bar{g}}[\hat{g}_S] \geq \mathscr{G}^{\mathcal{S}}_{M,\bar{g}}[\bar{g}],$$

we have $\hat{g}_S \in \mathcal{Q}_{\bar{g}}$. That is, $\hat{g}_S = c^2\bar{g}$ for some constant $c > 0$. □

Now we can prove the volume comparison of Einstein manifolds with respect to scalar curvature.

Proof of Theorem B. According to Theorem 5.6, we can find a local slice $\mathcal{S}_{\bar{g}}$ through the reference metric $\bar{g}$. Moreover, there exists a constant $\varepsilon_0 > 0$ such that for any metric $\tilde{g}$ with

$$\|\tilde{g} - \bar{g}\|_{C^2(M,\bar{g})} < \varepsilon_0,$$

we can find a diffeomorphism $\psi \in \mathscr{D}(M)$ with the property that $\psi^*\tilde{g} \in U_{\bar{g}} \subseteq \mathcal{S}_{\bar{g}}$, where the subset $U_{\bar{g}}$ is given by Proposition 5.8.

For $\lambda \neq 0$, suppose g is a metric on M with scalar curvature

$$R_g \geq n(n-1)\lambda$$

and

$$\|g - \bar{g}\|_{C^2(M,\bar{g})} < \varepsilon_0.$$

In addition, we assume the reverse inequality of the claimed volume comparison:

$$\lambda(V_M(g) - V_M(\bar{g})) \geq 0. \tag{5-5}$$

This implies there is a diffeomorphism $\varphi \in \mathscr{D}(M)$ such that $\varphi^* g \in U_{\bar{g}} \subseteq \mathcal{S}_{\bar{g}}$ and

$$\mathscr{G}^{\mathcal{S}}_{M,\bar{g}}[\varphi^* g] = V_M(\varphi^* g)^{2/n}\int_M (R_g \circ \varphi)\, dv_{\bar{g}} \geq V_M(\bar{g})^{2/n}\int_M R_{\bar{g}}\, dv_{\bar{g}} = \mathscr{G}^{\mathcal{S}}_{M,\bar{g}}[\bar{g}],$$

due to our assumptions and the fact that $R_{\bar{g}} = n(n-1)\lambda$ is a constant. According to Proposition 5.8, there exists a constant $c > 0$ such that $\varphi^* g = c^2\bar{g}$.

From our assumptions,

$$R_{\varphi^* g} = c^{-2} R_{\bar{g}} \geq R_{\bar{g}} = n(n-1)\lambda,$$

and hence
$$\lambda(1-c) \geq 0.$$
However, inequality (5-5) suggests that
$$0 \leq \lambda(V_M(\varphi^* g) - V_M(\bar{g})) = \lambda(c^n - 1) V_M(\bar{g}),$$
which implies that $\lambda(1-c) \leq 0$. Therefore, we conclude $c = 1$ and hence $\varphi^* g = \bar{g}$. That is, (M^n, g) is isometric to $(M^n, \bar{g})$, and this concludes the theorem. □

With analogous techniques, we can prove the local rigidity of Ricci-flat manifolds.

Proof of Theorem C. Similar to the proof of Theorem B, we can find a constant $\varepsilon_0 > 0$ such that for any metric $\tilde{g}$ satisfying
$$\|\tilde{g} - \bar{g}\|_{C^2(M,\bar{g})} < \varepsilon_0,$$
there exists a diffeomorphism $\varphi \in \mathscr{D}(M)$ such that $\varphi^* g \in U_{\bar{g}} \subseteq \mathcal{S}_{\bar{g}}$, where $U_{\bar{g}}$ is given in Proposition 5.8.

Suppose g is a Riemannian metric with scalar curvature
$$R_g \geq 0$$
and
$$\|g - \bar{g}\|_{C^2(M,\bar{g})} < \varepsilon_0.$$
Then there is a diffeomorphism $\varphi \in \mathscr{D}(M)$ such that
$$\mathscr{G}^{\mathcal{S}}_{M,\bar{g}}[\varphi^* g] = V_M(\varphi^* g)^{2/n} \int_M (R_g \circ \varphi)\, dv_{\bar{g}} \geq 0.$$
However,
$$\mathscr{G}^{\mathcal{S}}_{M,\bar{g}}[\bar{g}] = V_M(\bar{g})^{2/n} \int_M R_{\bar{g}}\, dv_{\bar{g}} = 0,$$
and hence there is a constant $c > 0$ such that $\varphi^* g = c^2 \bar{g}$ due to $\bar{g}$ being strictly stable Ricci-flat and Proposition 5.8. The conclusion follows. □

According to Proposition 5.3, the second variation of $\mathscr{G}_{M,\bar{g}}$ at an unstable Einstein metric $\bar{g}$ is indefinite and hence $\bar{g}$ is a saddle point instead of a local maximum. This suggests that the volume comparison may fail for unstable Einstein manifolds and counterexamples can be constructed. It is well known that a product of positive Einstein manifolds with identical Einstein constants is still Einstein but unstable; see [Kröncke 2013]. Due to this reason and its simple structure, it can be our first choice.

The following example is constructed by Macbeth (personal communication, 2019), which shows the stability assumption is necessary for our volume comparison theorem.

Proposition 5.9. *There is a family of metrics $\{g_t\}_{t\in[0,1)}$ on $\mathbb{S}^2 \times \mathbb{S}^2$ such that*

- *g_0 is the canonical product metric on $\mathbb{S}^2 \times \mathbb{S}^2$,*
- *$R_{g_t} = R_{g_{\mathbb{S}^2\times\mathbb{S}^2}} = 4$ for all $t \in [0, 1)$,*
- *$V_M(g_t) > V_M(g_{\mathbb{S}^2\times\mathbb{S}^2})$ for all $t \in (0, 1)$.*

Proof. Let
$$g_t = (1+t)^{-1} g^1_{\mathbb{S}^2} + (1-t)^{-1} g^2_{\mathbb{S}^2}$$
with $t \in [0, 1)$, where $g^i_{\mathbb{S}^2}$ is the canonical metric on the i-th $\mathbb{S}^2$ factor, $i = 1, 2$. It is easy to see that their scalar curvature is given by
$$R_{g_t} = 2(1+t) + 2(1-t) = 4$$
for all $t \in [0, 1)$. However, its volume is
$$V_{\mathbb{S}^2\times\mathbb{S}^2}(g_t) = (1-t^2)^{-1} V_{\mathbb{S}^2\times\mathbb{S}^2}(\bar{g}) > V_{\mathbb{S}^2\times\mathbb{S}^2}(\bar{g}). \qquad \square$$

It is straightforward to generalize this example to more general product cases. It would be interesting to see whether we can find an explicit example of an unstable Einstein manifold which is not of this type but where the volume comparison fails.

Appendix: Equivalence of Schoen's conjectures

In this appendix, we show that two well-known conjectures proposed by Schoen [1989] on hyperbolic manifolds actually are equivalent to each other. We believe the proof is known to experts. Unfortunately, we could not find an appropriate reference. Thus we present a proof here for interested readers.

We start with a well-known concept in conformal geometry; see [Viaclovsky 2016].

Definition A.1. For $n \geq 3$, let (M^n, g) be a connected closed n-dimensional Riemannian manifold. The *Yamabe constant* of the conformal class $[g]$ is defined to be
$$Y(M^n, [g]) := \inf_{g \in [g]} \frac{\int_M R_g \, dv_g}{(V_M(g))^{(n-2)/n}}.$$
Moreover, we can define a min-max invariant
$$Y(M^n) := \sup_{[g]} Y(M^n, [g])$$
called the *Yamabe invariant* or *σ-invariant*.

It is well known that
$$Y(M^n) \leq Y(\mathbb{S}^n)$$
for any closed smooth manifold M^n and the canonical spherical metric achieves the Yamabe invariant of $\mathbb{S}^n$. For a given closed hyperbolic manifold with dimension at least three, its hyperbolic metric is unique up to a dilation due to the well-known *Mostow rigidity theorem*; see Theorem C.0 in [Benedetti and Petronio 1992]. Similar to the spherical case, Schoen [1989] conjectures that its Yamabe invariant is achieved by the canonical hyperbolic metric.

Conjecture A (Schoen's hyperbolic Yamabe invariant conjecture). *For $n \geq 3$, suppose $(M^n, \bar{g})$ is an n-dimensional closed hyperbolic manifold. Then*
$$Y(M^n) = Y(M^n, [\bar{g}]),$$
i.e., the Yamabe invariant is achieved by its canonical hyperbolic metric.

Another conjecture about closed hyperbolic manifolds concerns volume comparison, which is also referred to as *Schoen's conjecture*.

Conjecture B (Schoen's hyperbolic volume comparison conjecture). *For $n \geq 3$, suppose $(M^n, \bar{g})$ is an n-dimensional closed hyperbolic manifold. Then for any metric g on M with scalar curvature*

$$R_g \geq R_{\bar{g}},$$

its volume satisfies

$$V_M(g) \geq V_M(\bar{g}).$$

Obviously, Conjecture A involves all metrics on the given hyperbolic manifold and in general it is difficult to solve. Conjecture B only involves the comparison of a special metric with the reference metric, which seems easier to solve than Conjecture A. However, Conjectures A and B are in fact equivalent to each other and hence they are equally difficult in this sense. The bright side of this equivalence is that we only need to solve Conjecture B, then Conjecture A will hold automatically. This seems to be a promising approach to Conjecture A.

In the rest of the appendix, we will show the equivalence of Conjectures A and B.

We first show Conjecture A implies Conjecture B. In order to do this, we need the following lemma adapted from an observation of Kobayashi [1987].

Lemma A.2. *Let (M^n, g) be a closed manifold and $Y(M^n, [g])$ be the Yamabe constant of the conformal class $[g]$. Then*

$$-\left(\int_M |R_g^-|^{n/2}\, dv_g\right)^{2/n} \leq Y(M^n, [g]) \leq \left(\int_M |R_g^+|^{n/2}\, dv_g\right)^{2/n},$$

where $R_g^+ := \max\{R_g, 0\}$ and $R_g^- := \max\{-R_g, 0\}$.

Proof. By the conformal transformation law of scalar curvature,

$$Y(M^n, [g]) = \inf_{u>0} \frac{\int_M (a|\nabla_g u|_g^2 + R_g u^2)\, dv_g}{\left(\int_M u^{2n/(n-2)}\, dv_g\right)^{(n-2)/n}},$$

where $a := 4(n-1)/(n-2)$. Then we have

$$Y(M^n, [g]) \geq \inf_{u>0} \frac{\int_M R_g u^2\, dv_g}{\left(\int_M u^{2n/(n-2)}\, dv_g\right)^{(n-2)/n}} \geq -\inf_{u>0} \frac{\int_M R_g^- u^2\, dv_g}{\left(\int_M u^{2n/(n-2)}\, dv_g\right)^{(n-2)/n}},$$

since $R_g = R_g^+ - R_g^-$. By *Hölder's inequality*,

$$\int_M R_g^- u^2\, dv_g \leq \left(\int_M |R_g^-|^{n/2}\, dv_g\right)^{2/n} \left(\int_M u^{2n/(n-2)}\, dv_g\right)^{(n-2)/n},$$

and hence

$$Y(M^n, [g]) \geq -\left(\int_M |R_g^-|^{n/2}\, dv_g\right)^{2/n}.$$

Similarly,

$$Y(M^n, [g]) \leq \frac{\int_M R_g dv_g}{(V_M(g))^{(n-2)/n}} \leq \frac{\int_M R_g^+ dv_g}{(V_M(g))^{(n-2)/n}}.$$

By *Hölder's inequality*,

$$\int_M R_g^+ dv_g \le \left(\int_M |R_g^+|^{n/2}\, dv_g\right)^{2/n} (V_M(g))^{(n-2)/n},$$

and hence

$$Y(M^n,[g]) \le \left(\int_M |R_g^+|^{n/2}\, dv_g\right)^{2/n}.$$

□

Immediately, this implies the following conformal volume comparison.

Proposition A.3. *Suppose $(M^n, \hat{g})$ is a closed Riemannian manifold with strictly negative constant scalar curvature $R_{\hat{g}}$. Then for any metric $g \in [\hat{g}]$ with scalar curvature*

$$R_g \ge R_{\hat{g}},$$

we have

$$V_M(g) \ge V_M(\hat{g}).$$

Proof. Since $R_{\hat{g}}$ is a strictly negative constant, then its Yamabe constant satisfies

$$Y(M^n,[\hat{g}]) < 0,$$

and hence $\hat{g}$ is a Yamabe metric in the conformal class $[\hat{g}]$ due to the uniqueness of the Yamabe metric of negative Yamabe constant. Thus,

$$Y(M^n,[\hat{g}]) = R_{\hat{g}}(V_M(\hat{g}))^{2/n}.$$

By Lemma A.2,

$$\left(\min_M R_g\right)(V_M(g))^{2/n} - \left(\int_M |R_g^-|^{n/2} dv_g\right)^{n/2} \le Y(M^n,[\hat{g}]) = R_{\hat{g}}(V_M(\hat{g}))^{2/n}.$$

Therefore,

$$R_{\hat{g}}(V_M(g))^{2/n} \le \left(\min_M R_g\right)(V_M(g))^{2/n} \le R_{\hat{g}}(V_M(\hat{g}))^{2/n},$$

and hence

$$V_M(g) \ge V_M(\hat{g}).$$

□

Proposition A.4. *Conjecture A* $\Longrightarrow$ *Conjecture B.*

Proof. Let $(M^n, \bar{g})$ be a closed hyperbolic manifold. Suppose g is a metric on M with scalar curvature

$$R_g \ge R_{\bar{g}}.$$

We are going to show

$$V_M(g) \ge V_M(\bar{g}),$$

assuming $\bar{g}$ achieves its Yamabe invariant $Y(M^n)$.

From Conjecture A, the Yamabe constant of the conformal class $[g]$ satisfies

$$Y(M^n,[g]) \le Y(M^n) = Y(M^n,[\bar{g}]) < 0.$$

Let $\hat{g} \in [g]$ be the unique Yamabe metric in $[g]$ which is normalized such that $R_{\hat{g}} = R_{\bar{g}}$. By Proposition A.3, we have

$$V_M(g) \geq V_M(\hat{g}).$$

On the other hand,

$$R_{\hat{g}} V_M(\hat{g})^{2/n} = Y(M^n, [g]) \leq Y(M^n) = Y(M^n, [\bar{g}]) = R_{\bar{g}} V_M(\bar{g})^{2/n},$$

which implies

$$V_M(\hat{g}) \geq V_M(\bar{g}).$$

Therefore

$$V_M(g) \geq V_M(\hat{g}) \geq V_M(\bar{g}),$$

and hence Conjecture B holds. □

Proposition A.5. *Conjecture B* $\Longrightarrow$ *Conjecture A.*

Proof. Let $(M^n, \bar{g})$ be a closed hyperbolic manifold. We will show that its Yamabe invariant satisfies

$$Y(M^n) = Y(M^n, [\bar{g}]),$$

assuming the volume comparison holds.

We first recall a classic result of Gromov and Lawson [1983, Corollary A] which states that there is no metric with nonnegative scalar curvature on a compact hyperbolic manifold. That means the Yamabe invariant satisfies

$$Y(M^n) \leq 0,$$

and there is no metric on M with identically vanishing scalar curvature. Thus for any metric g on M, the Yamabe constant of the conformal class $[g]$ is strictly negative:

$$Y(M^n, [g]) < 0.$$

Let $\hat{g}$ be the Yamabe metric in the conformal class $[g]$ with $R_{\hat{g}} = R_{\bar{g}} < 0$. According to Conjecture B,

$$V_M(\hat{g}) \geq V_M(\bar{g}).$$

Therefore, the Yamabe constant of $[g]$ satisfies

$$Y(M^n, [g]) = \frac{\int_M R_{\hat{g}}\, dv_{\hat{g}}}{(V_M(\hat{g}))^{(n-2)/2}} = R_{\hat{g}}(V_M(\hat{g}))^{2/n} \leq R_{\bar{g}}(V_M(\bar{g}))^{2/n} = Y(M^n, [\bar{g}]).$$

Since g is arbitrary, we conclude

$$Y(M^n) = \sup_{[g]} Y(M^n, [g]) = Y(M^n, [\bar{g}]),$$

and hence Conjecture A holds. □

In summary, we have the equivalence of Schoen's Conjectures A and B.

Theorem A.6. *Conjecture A* $\Longleftrightarrow$ *Conjecture B.*

Acknowledgements

The author would like to express his appreciation to Professor Miao Pengzi for suggesting this interesting problem. The author would also like to thank Professors Chen Bing-Long, Lin Yueh-Ju, Heather Macbeth, Qing Jie, Wei Guofang, Zhang Hui-Chun and Zhu Xi-Ping for their interest in this problem and inspiring discussions. In particular, the author would like to express his appreciation to Professor Zhang Hui-Chun for pointing out the work [Besson et al. 1991], Professor Zhu Xi-Ping for explaining the idea of the work [Agol et al. 2007] and also many other valuable comments, Professor Heather Macbeth for sharing the counterexample for unstable Einstein manifolds and suggestions for considering the stability of Einstein metrics, and Professor Wei Guofang for valuable comments. The author would also like to thank the referee for many valuable comments and suggestions on this article.

References

[Agol et al. 2007] I. Agol, P. A. Storm, and W. P. Thurston, "Lower bounds on volumes of hyperbolic Haken 3-manifolds", *J. Amer. Math. Soc.* **20**:4 (2007), 1053–1077. MR Zbl

[Baltazar and Ribeiro 2017] H. Baltazar and E. Ribeiro, Jr., "Critical metrics of the volume functional on manifolds with boundary", *Proc. Amer. Math. Soc.* **145**:8 (2017), 3513–3523. MR Zbl

[Barros et al. 2015] A. Barros, R. Diógenes, and E. Ribeiro, Jr., "Bach-flat critical metrics of the volume functional on 4-dimensional manifolds with boundary", *J. Geom. Anal.* **25**:4 (2015), 2698–2715. MR Zbl

[Benedetti and Petronio 1992] R. Benedetti and C. Petronio, *Lectures on hyperbolic geometry*, Springer, 1992. MR Zbl

[Besse 1987] A. L. Besse, *Einstein manifolds*, Ergebnisse der Mathematik (3) **10**, Springer, 1987. MR Zbl

[Besson et al. 1991] G. Besson, G. Courtois, and S. Gallot, "Volume et entropie minimale des espaces localement symétriques", *Invent. Math.* **103**:2 (1991), 417–445. MR Zbl

[Besson et al. 1995] G. Besson, G. Courtois, and S. Gallot, "Entropies et rigidités des espaces localement symétriques de courbure strictement négative", *Geom. Funct. Anal.* **5**:5 (1995), 731–799. MR Zbl

[Bray 1997] H. L. Bray, *The Penrose inequality in general relativity and volume comparison theorems involving scalar curvature*, Ph.D. thesis, Stanford University, 1997, available at https://www.proquest.com/docview/304386501.

[Brendle 2012] S. Brendle, "Rigidity phenomena involving scalar curvature", pp. 179–202 in *Surveys in differential geometry, XVII* (Bethlehem, PA, 2010), edited by H.-D. Cao and S.-T. Yau, International, Boston, 2012. MR Zbl

[Brendle and Marques 2011] S. Brendle and F. C. Marques, "Scalar curvature rigidity of geodesic balls in $\mathbb{S}^n$", *J. Differential Geom.* **88**:3 (2011), 379–394. MR Zbl

[Chow et al. 2006] B. Chow, P. Lu, and L. Ni, *Hamilton's Ricci flow*, Grad. Stud. Math. **77**, Amer. Math. Soc., Providence, RI, 2006. MR Zbl

[Corvino et al. 2013] J. Corvino, M. Eichmair, and P. Miao, "Deformation of scalar curvature and volume", *Math. Ann.* **357**:2 (2013), 551–584. MR Zbl

[Dai et al. 2005] X. Dai, X. Wang, and G. Wei, "On the stability of Riemannian manifold with parallel spinors", *Invent. Math.* **161**:1 (2005), 151–176. MR Zbl

[Dai et al. 2007] X. Dai, X. Wang, and G. Wei, "On the variational stability of Kähler–Einstein metrics", *Comm. Anal. Geom.* **15**:4 (2007), 669–693. MR Zbl

[Ebin 1970] D. G. Ebin, "The manifold of Riemannian metrics", pp. 11–40 in *Global analysis* (Berkeley, CA, 1968), edited by S.-s. Chern and S. Smale, Proc. Symp. Pure Math. **15**, Amer. Math. Soc., Providence, RI, 1970. MR Zbl

[Fischer and Marsden 1975] A. E. Fischer and J. E. Marsden, "Deformations of the scalar curvature", *Duke Math. J.* **42**:3 (1975), 519–547. MR Zbl

[Gromov and Lawson 1980] M. Gromov and H. B. Lawson, Jr., "Spin and scalar curvature in the presence of a fundamental group, I", *Ann. of Math.* (2) **111**:2 (1980), 209–230. MR Zbl

[Gromov and Lawson 1983] M. Gromov and H. B. Lawson, Jr., "Positive scalar curvature and the Dirac operator on complete Riemannian manifolds", *Inst. Hautes Études Sci. Publ. Math.* **58** (1983), 83–196. MR Zbl

[Gursky and Viaclovsky 2004] M. J. Gursky and J. A. Viaclovsky, "Volume comparison and the σ_k-Yamabe problem", *Adv. Math.* **187**:2 (2004), 447–487. MR Zbl

[Hamilton 1999] R. S. Hamilton, "Non-singular solutions of the Ricci flow on three-manifolds", *Comm. Anal. Geom.* **7**:4 (1999), 695–729. MR Zbl

[Kobayashi 1987] O. Kobayashi, "Scalar curvature of a metric with unit volume", *Math. Ann.* **279**:2 (1987), 253–265. MR Zbl

[Koiso 1980] N. Koiso, "Rigidity and stability of Einstein metrics: the case of compact symmetric spaces", *Osaka Math. J.* **17**:1 (1980), 51–73. MR Zbl

[Kröncke 2013] K. Kröncke, *Stability of Einstein manifolds*, Ph.D. thesis, University of Potsdam, 2013, available at http://opus.kobv.de/ubp/volltexte/2014/6963.

[Li 2012] P. Li, *Geometric analysis*, Cambridge Stud. Adv. Math. **134**, Cambridge Univ. Press, 2012. MR Zbl

[Lin and Yuan 2022] Y.-J. Lin and W. Yuan, "Deformations of Q-curvature, II", *Calc. Var. Partial Differential Equations* **61**:2 (2022), art. id 74, 28 pp. MR

[Miao and Tam 2009] P. Miao and L.-F. Tam, "On the volume functional of compact manifolds with boundary with constant scalar curvature", *Calc. Var. Partial Differential Equations* **36**:2 (2009), 141–171. MR Zbl

[Miao and Tam 2012] P. Miao and L.-F. Tam, "Scalar curvature rigidity with a volume constraint", *Comm. Anal. Geom.* **20**:1 (2012), 1–30. MR Zbl

[Perelman 2002] G. Perelman, "The entropy formula for the Ricci flow and its geometric applications", preprint, 2002. Zbl arXiv math/0211159

[Perelman 2003] G. Perelman, "Ricci flow with surgery on three-manifolds", preprint, 2003. Zbl arXiv math/0303109

[Qing and Yuan 2013] J. Qing and W. Yuan, "A note on static spaces and related problems", *J. Geom. Phys.* **74** (2013), 18–27. MR Zbl

[Qing and Yuan 2016] J. Qing and W. Yuan, "On scalar curvature rigidity of vacuum static spaces", *Math. Ann.* **365**:3-4 (2016), 1257–1277. MR Zbl

[Rosenberg 2007] J. Rosenberg, "Manifolds of positive scalar curvature: a progress report", pp. 259–294 in *Surveys in differential geometry, XI: Metric and comparison theory*, edited by J. Cheeger and K. Grove, International, 2007. MR Zbl

[Schoen 1989] R. M. Schoen, "Variational theory for the total scalar curvature functional for Riemannian metrics and related topics", pp. 120–154 in *Topics in calculus of variations* (Montecatini Terme, Italy, 1987), edited by M. Giaquinta, Lecture Notes in Math. **1365**, Springer, 1989. MR Zbl

[Schoen and Yau 1979a] R. Schoen and S. T. Yau, "Existence of incompressible minimal surfaces and the topology of three-dimensional manifolds with non-negative scalar curvature", *Ann. of Math.* (2) **110**:1 (1979), 127–142. MR Zbl

[Schoen and Yau 1979b] R. Schoen and S. T. Yau, "On the structure of manifolds with positive scalar curvature", *Manuscripta Math.* **28**:1-3 (1979), 159–183. MR Zbl

[Viaclovsky 2016] J. A. Viaclovsky, "Critical metrics for Riemannian curvature functionals", pp. 197–274 in *Geometric analysis* (Park City, UT, 2013), edited by H. L. Bray et al., IAS/Park City Math. Ser. **22**, Amer. Math. Soc., Providence, RI, 2016. MR Zbl

[Yuan 2015] W. Yuan, *The geometry of vacuum static spaces and deformations of scalar curvature*, Ph.D. thesis, University of California, Santa Cruz, 2015, available at https://www.proquest.com/docview/1710736908.

[Zhang 2019] Y. Zhang, "Scalar curvature volume comparison theorems for almost rigid sphere", preprint, 2019. arXiv 1909.00909

Received 3 Jun 2017. Revised 18 Feb 2021. Accepted 4 May 2021.

WEI YUAN: yuanw9@mail.sysu.edu.cn
Department of Mathematics, Sun Yat-sen University, Guangzhou, China

ANALYSIS AND PDE
Vol. 16 (2023), No. 1, pp. 35–88
DOI: 10.2140/apde.2023.16.35

WANDERING DOMAINS ARISING FROM LAVAURS MAPS WITH SIEGEL DISKS

MATTHIEU ASTORG, LUKA BOC THALER AND HAN PETERS

The first example of polynomial maps with wandering domains was constructed in 2016 by the first and last authors, together with Buff, Dujardin and Raissy. In this paper, we construct a second example with different dynamics, using a Lavaurs map with a Siegel disk instead of an attracting fixed point. We prove a general necessary and sufficient condition for the existence of a trapping domain for nonautonomous compositions of maps converging parabolically towards a Siegel-type limit map. Constructing a skew-product satisfying this condition requires precise estimates on the convergence to the Lavaurs map, which we obtain by a new approach. We also give a self-contained construction of parabolic curves, which are integral to this new method.

1. Introduction

Rational functions do not have wandering domains, a classical result due to [Sullivan 1985]. Recently in [Astorg et al. 2016] it was shown that there do exist polynomial maps in two complex variables with wandering Fatou components. The maps constructed in [Astorg et al. 2016] are polynomial skew products of the form

$$(z, w) \mapsto (f_w(z), g(w)),$$

where $g(w)$ and $f_w(z) = f(z, w)$ are polynomials in respectively one and two variables. While the construction holds for families of maps with arbitrarily many parameters, the constructed examples are essentially unique: they all arise from similar behavior and cannot easily be distinguished in terms of the geometry of the components or qualitative behavior of the orbits in the components. The goal in this paper is to modify the construction in [Astorg et al. 2016] to obtain quite different examples of wandering Fatou components. Our construction requires much more precise convergence estimates, forcing us to revisit and clarify the original proof, obtaining a better understanding of the methodology.

The maps considered in [Astorg et al. 2016] are of the specific form

$$P : (z, w) \mapsto \left(f(z) + \tfrac{\pi^2}{4} w, g(w)\right), \tag{1}$$

where $f(z) = z + z^2 + O(z^3)$ and $g(w) = w - w^2 + O(w^3)$. Recall that the constant $\frac{\pi^2}{4}$ is essential to guarantee the following key result in [Astorg et al. 2016]:

This research was partially supported by the ANR grant Fatou ANR-17-CE40-0002-01. Boc Thaler was supported by the SIR grant “NEWHOLITE - New methods in holomorphic iteration” no. RBSI14CFME and by the research program P1-0291 from ARRS, Republic of Slovenia.
MSC2010: 32H50, 37F10, 37G15.
Keywords: Fatou set, holomorphic dynamics, parabolic implosion, polynomial mappings, skew-products, wandering Fatou components, parabolic curves, nonautonomous dynamics.

Proposition A. *As $n \to +\infty$, the sequence of maps*

$$(z, w) \mapsto P^{\circ 2n+1}(z, g^{\circ n^2}(w))$$

converges locally uniformly in $\mathcal{B}_f \times \mathcal{B}_g$ to the map

$$(z, w) \mapsto (\mathcal{L}_f(z), 0).$$

Throughout this paper $\mathcal{B}_f$ and $\mathcal{B}_g$ refer to the *parabolic basins* of, respectively, f and g, and $\mathcal{L}_f$ refers to the *Lavaurs map* of f with phase 0; see for example [Lavaurs 1989; Shishikura 2000]. By carefully choosing the higher-order terms of f, one can select Lavaurs maps with desired dynamical behavior.

In Proposition B of [Astorg et al. 2016] it was shown that $\mathcal{L}_f$ can have an attracting fixed point. The fact that P has a wandering Fatou component is then a quick corollary of Proposition A. It seems very likely that one can similarly construct wandering domains when $\mathcal{L}_f$ has a parabolic fixed point, using the refinement of Proposition A presented here.

We will construct wandering domains arising when $\mathcal{L}_f$ has a Siegel fixed point: an irrationally indifferent fixed point with Diophantine rotation number. Compositions of small perturbations of $\mathcal{L}_f$ behave so subtly that it is far from clear that Lavaurs maps with Siegel disks can produce wandering domains.

In order to control the behavior of successive perturbations, we prove a refinement of Proposition A with precise convergence estimates, showing that the convergence towards the Lavaurs map is "parabolic". Moreover, we study the behavior of nonautonomous systems given by maps converging parabolically to a limit map with a Siegel fixed point. We introduce an easily computable index characterizing the behavior of the nonautonomous systems.

In the next section we give more precise statements of our results, and prove how the combination of these results provides a new construction of wandering domains.

2. Background and overview of results

2A. ***Polynomial skew products and Fatou components.*** There is more than one possible interpretation of Fatou and Julia sets for polynomial skew products; see for example [Jonsson 1999] for a thorough discussion. When we discuss Fatou components of skew products here, we consider open connected sets in $\mathbb{C}^2$ whose orbits are uniformly bounded, which of course implies equicontinuity. Since the degrees of f and g in (1) are at least 2, the complement of a sufficiently large bidisk is contained in the escape locus, which is connected; all other Fatou components are therefore bounded and have bounded orbits.

Given a Fatou component U of P, normality implies that its projection onto the second coordinate $\pi_w(U)$ is contained in a Fatou component of g, which must therefore be periodic or preperiodic. Without loss of generality we may assume that this component of g is invariant, and thus either an attracting basin, a parabolic basin or a Siegel disk.

The behavior of P inside a Siegel disk of g may be very complicated and has received little attention in the literature, but see [Peters and Raissy 2019] for the treatment of a special case.

There have been a number of results proving the nonexistence of wandering domains inside attracting basins of g. The nonexistence of wandering domains in the superattracting case was proved in [Lilov

2004], but it was shown in [Peters and Vivas 2016] that the arguments from Lilov cannot hold in the geometrically attracting case. The nonexistence of wandering domains under progressively weaker conditions was proved in [Peters and Smit 2018; Ji 2020].

Here, as in [Astorg et al. 2016], we will consider components U for which $\pi_w(U)$ is contained in a parabolic basin of g. We assume that the fixed point of g lies at the origin, and that g is of the form $g(w) = w - w^2 + \text{h.o.t.}$, so that orbits approach 0 tangent to the positive real axis. We will in fact make the stronger assumption $g(w) = w - w^2 + w^3 + \text{h.o.t.}$

2B. ***Fatou coordinates and Lavaurs' theorem.*** Consider a polynomial $f(z) = z - z^2 + az^3 + \text{h.o.t.}$ For $r > 0$ small enough we define incoming and outgoing petals

$$P_f^\iota = \{|z + r| < r\} \quad \text{and} \quad P_f^o = \{|z - r| < r\}.$$

The incoming petal P_f^ι is forward invariant, and all orbits in P_f^ι converge to 0. Moreover, any orbit which converges to 0 but never lands at 0 must eventually be contained in P_f^ι. Therefore we can define the parabolic basin as

$$\mathcal{B}_f = \bigcup f^{-n} P_f^\iota.$$

The outgoing petal P_f^o is backwards invariant, with backwards orbits converging to 0.

On P_f^ι and P_f^o one can define incoming and outgoing Fatou coordinates $\phi_f^\iota : P_f^\iota \to \mathbb{C}$ and $\phi_f^o : P_f^o \to \mathbb{C}$ solving the functional equations

$$\phi_f^\iota \circ f(z) = \phi_f^\iota(z) + 1 \quad \text{and} \quad \phi_f^o \circ f(z) = \phi_f^o(z) + 1,$$

where $\phi_f^\iota(P_f^\iota)$ contains a right half-plane and $\phi_f^o(P_f^o)$ contains a left half-plane. By the first functional equation the incoming Fatou coordinates can be uniquely extended to the attracting basin $\mathcal{B}_f$. On the other hand, the inverse of ϕ_f^o, denoted by ψ_f^o, can be extended to the entire complex plane, still satisfying the functional equation

$$f \circ \psi_f^o(Z) = \psi_f^o(z + 1).$$

The fact that the exceptional set of f is empty implies that $\psi_f^o : \mathbb{C} \to \mathbb{C}$ is surjective. We note that both incoming and outgoing Fatou coordinates are (on the corresponding petals) of the form $Z = -1/z + b\log(z) + o(1)$, where the coefficient b vanishes when $a = 1$. This is one reason for working with maps f of the form $f(z) = z + z^2 + z^3 + \text{h.o.t.}$

Let us now consider small perturbations of the map f. For $\epsilon \in \mathbb{C}$ we write $f_\epsilon(z) = f(z) + \epsilon^2$, and consider the behavior as $\epsilon \to 0$. The most interesting behavior occurs when ϵ approaches 0 tangent to the positive real axis.

Lavaurs' theorem [1989]. *Let $\epsilon_j \to 0$, $n_j \in \mathbb{N}$ and $\alpha \in \mathbb{C}$ satisfy*

$$n_j - \frac{\pi}{\epsilon_j} \to \alpha \quad \textit{as } j \to \infty.$$

Then

$$f_{\epsilon_j}^{n_j} \to \mathcal{L}_f(\alpha) = \psi_f^o \circ \tau_\alpha \circ \phi_f^\iota,$$

where $\tau_\alpha(Z) = Z + \alpha$.

The map $\mathcal{L}_f(\alpha)$ is called the *Lavaurs map*, and α is called the *phase*. In this paper we will only consider phase $\alpha = 0$, and write $\mathcal{L}_f$ instead of $\mathcal{L}_f(0)$.

2C. *Propositions A and B.* The construction of wandering domains in [Astorg et al. 2016] follows quickly from two key propositions, the aforementioned Propositions A and B. In this paper we will prove a variation to Proposition B, and a refinement to Proposition A, which we will both state here.

Our main technical result is the following refinement of Proposition A. As before we write $P(z,w) = \big(f(z) + \frac{\pi^2}{4}w, g(w)\big)$, with $f(z) = z + z^2 + z^3 + bz^4 + \text{h.o.t.}$, and $g(w) = w - w^2 + w^3 + \text{h.o.t.}$

Proposition A′. *There exists a holomorphic function $h : \mathcal{B}_f \times \mathcal{B}_g \to \mathbb{C}$ such that*

$$P^{2n+1}(z, g^{n^2}(w)) = (\mathcal{L}_f(z), 0) + \left(\frac{h(z,w)}{n}, 0\right) + O\left(\frac{\log n}{n^2}\right),$$

uniformly on compact subsets of $\mathcal{B}_f \times \mathcal{B}_g$. The function $h(z,w)$ is given by

$$h(z,w) = \frac{\mathcal{L}_f'(z)}{(\phi_f^\iota)'(z)} \cdot (C + \phi_f^\iota(z) - \phi_g^\iota(w)),$$

where the constant $C \in \mathbb{C}$ depends on b.

Proposition A′ will be proved in Section 5; see Theorem 5.33.

Proposition B in [Astorg et al. 2016] states that the Lavaurs map $\mathcal{L}_f$ of a polynomial $f(z) = z + z^2 + az^3 + O(z^4)$ has an attracting fixed point for suitable choices of the constant $a \in \mathbb{C}$. We recall very briefly the main idea in the proof of Proposition B: For $a = 1$ the “horn map” has a parabolic fixed point at infinity. By perturbing $a \simeq 1$, the parabolic fixed point bifurcates, and for appropriate perturbations this guarantees the existence of an attracting fixed point for the horn map, and thus also for the Lavaurs map.

In this paper we will consider a more restrictive family of polynomials of the form $f(z) = z + z^2 + z^3 + O(z^4)$, which means that we cannot use the above bifurcation argument. Using a different line of reasoning, using small perturbations of a suitably chosen degree-7 real polynomial, we will prove a variation to Proposition B, namely Proposition B′ below. The proof of Proposition B′ will be given in Section 6.

Before stating the proposition we recall that a fixed point $z_0 = \mathcal{L}_f(z_0)$ is said to be of Siegel type if $\lambda = \mathcal{L}_f'(z_0) = e^{2\pi i \zeta}$, where $\zeta \in \mathbb{R} \backslash \mathbb{Q}$ is Diophantine, i.e., if there exist $c, r > 0$ such that $|\lambda^n - 1| \geq cn^{-r}$ for all integers $n > 0$. Recall that neutral fixed points with Diophantine rotation numbers are always locally linearizable:

Theorem 2.1 [Siegel 1942]. *Let $p(z) = e^{2\pi i \zeta}z + O(z^2)$ be a holomorphic germ. If ζ is Diophantine then there exist a neighborhood of the origin Ω_p and a biholomorphic map $\varphi : \Omega_p \to D_r(0)$ of the form $\varphi(z) = z + a_2 z^2 + O(z^3)$ satisfying*

$$\varphi(p(z)) = e^{2\pi i \zeta}\varphi(z).$$

Proposition B′. *There exist polynomials of the form $f(z) = z + z^2 + z^3 + O(z^4)$ for which the Lavaurs map $\mathcal{L}_f$ has a* ***Siegel fixed point*** *z_0, with $\lambda = \mathcal{L}'_f(z_0)$. Moreover we can guarantee that*

$$\frac{\mathcal{L}''_f(z_0)(\phi^\iota_f)'(z_0)}{\lambda(1-\lambda)} - (\phi^\iota_f)''(z_0) \neq 0. \tag{2}$$

Condition (2) is necessary to guarantee the existence of wandering domains; see the discussion of the index κ later in this section, and the discussion in Section 5C.

A more precise description of the derivatives λ for which p is locally linearizable was given in [Bruno 1971; 1972; Yoccoz 1995]. As we are only concerned with constructing examples of maps with wandering Fatou components, we find it convenient to work with the stronger Diophantine condition. Proposition B′ will be proved in Section 6.

2D. *Perturbations of Siegel disks.* A key element in our study is the following question:

Let $f_1, f_2, \ldots$ be a sequence of holomorphic germs, converging locally uniformly to a holomorphic function f having a Siegel fixed point at 0. *Under which conditions does there exist a trapping region*?

By a trapping region we mean the existence of arbitrarily small neighborhoods U, V of 0 and $n_0 \in \mathbb{N}$ such that

$$f_m \circ \cdots \circ f_n(z) \in V$$

for all $z \in U$ and $m \geq n \geq n_0$. In other words, any orbit $(z_n)_{n\geq 0}$ that intersects U for sufficiently large n will afterwards be contained in a small neighborhood of the origin. Note that this in particular guarantees normality of the sequence of compositions $f_m \circ \cdots \circ f_0$ in a neighborhood of z_0, which is the reason for our interest in trapping regions.

We are particularly interested in the case where the differences $f_n - f$ are not absolutely summable, i.e., when

$$\sum_{n\geq n_0} \|f_n - f\|_U = \infty$$

for any n_0 and U. In this situation one generally does not expect a trapping region. However, motivated by Proposition A′, we will assume that $f_n - f$ is roughly of size $1/n$, and converges to zero along some real direction. More precisely, we assume that

$$f_n(z) - f(z) = \frac{h(z)}{n} + O\left(\frac{1}{z^{1+\epsilon}}\right), \tag{3}$$

where h is a holomorphic germ, defined in a neighborhood of the origin.

Theorem 2.2. *There exists an* ***index*** *κ, a rational expression in the coefficients of f and h, such that the following hold*:

(1) *If* $\mathrm{Re}(\kappa) = 0$, *then there is a trapping region, and all limit maps have rank* 1.

(2) *If* $\mathrm{Re}(\kappa) < 0$, *then there is a trapping region, and all orbits converge uniformly to the origin.*

(3) *If* $\mathrm{Re}(\kappa) > 0$, *then there is no trapping region. In fact, there can be at most one orbit that remains in a sufficiently small neighborhood of the origin.*

Theorem 2.2 holds under more general assumptions regarding the convergence towards the limit map, but the above statement is sufficient for our purposes. An example of a more general statement is given in Remark 3.15. An explicit formula for the index κ is given in Section 3, which contains the proof of Theorem 2.2.

Remark 2.3. The nonautonomous dynamics of the functions f_n satisfying (3) is closely related to the autonomous dynamics of the quasiparabolic map

$$F(z,w) = (f(z) + wh(z) + O(w^2), w - w^2 + O(w^2)).$$

The case $\mathrm{Re}(\kappa) < 0$ in Theorem 2.2 corresponds to F being dynamically separating and parabolically attracting, using the terminology of [Bracci and Zaitsev 2013]; by Corollary 6.3 of that work the map F has a connected basin of attraction at the origin. In particular this implies the existence of a trapping region for the sequence (f_n).

2E. *Parabolic curves.* An important idea in the proof of Lavaurs' theorem is that in a sufficiently small neighborhood of the origin, the function $f_\epsilon = f + \epsilon^2$ can be interpreted as a near-translation in the "almost Fatou coordinates": functions that converge to the ingoing and outgoing Fatou coordinates as $\epsilon \to \infty$. This idea is especially apparent in the treatment given in [Bedford et al. 2017]. The almost Fatou coordinates are defined using the pair of fixed points $\zeta_\pm(\epsilon)$ "splitting" from the parabolic fixed point.

When iterating two-dimensional skew products $P(z,w) = (f_w(z), g(w))$ it does not make sense to base the almost Fatou coordinates on the pair of fixed points of the maps $f_w(z) = f(z) + \frac{\pi^2}{4}w$, as the parameter w changes after every iteration of P. Instead, the natural idea would be to base these coordinates on a pair of invariant curves $\{z = \zeta_\pm(w)\}$, so-called *parabolic curves*, defined over a forward-invariant parabolic petal in the w-plane. The invariance of these parabolic curves is equivalent to the functional equations

$$\zeta_\pm(g(w)) = f_w(\zeta_\pm(w)).$$

In [Astorg et al. 2016], it is asked whether such parabolic curves exists. Instead, in that work it was shown that there exist *almost parabolic curves*, approximate solutions to the above functional equation with explicit error estimates. The proof of Proposition A relies to a great extent on these almost parabolic curves, and the fact that these are not exact solutions causes significant extra work.

In [López-Hernanz and Rosas 2020] it is shown that the parabolic curves indeed exist, in fact, the authors prove the existence of parabolic curves for any characteristic direction for diffeomorphisms in two complex dimensions. However, to be used in the proof of Proposition A, it is necessary to also obtain control over the domain of definition of the two parabolic curves. The result from [López-Hernanz and Rosas 2020] does not give the needed control.

In Section 4, Proposition 4.1, we give an alternative proof of the existence of parabolic curves, with control over the domains of definition. The availability of these parabolic curves forms an important ingredient in the proof of Proposition A$'$. The method of proof is a variation to the well-known graph transform method, and can likely be used to prove the existence of parabolic curves in greater generality.

2F. ***Wandering domains.*** Let us conclude this section by proving how Propositions A′ and B′ together imply the existence of wandering Fatou components. As before we let

$$P(z,w) = \left(f(z) + \tfrac{\pi^2}{4}w,\, g(w)\right),$$

where $g(w) = w - w^2 + w^3 + \text{h.o.t.}$ and the function $f(z) = z + z^2 + z^3 + \text{h.o.t.}$ is chosen such that $\mathcal{L}_f$ has a neutral fixed point z_0 with Diophantine rotation number. The existence of such f is given by Proposition B′.

Proposition A′ states that

$$P^{2n+1}(z, g^{n^2}(w)) = (\mathcal{L}_f(z), 0) + \left(\frac{h(z,w)}{n}, 0\right) + O\left(\frac{\log n}{n^2}\right),$$

uniformly on compact subsets of $\mathcal{B}_f \times \mathcal{B}_g$.

Recall from Proposition A′ that the function $h(z,w)$ is given by

$$h(z,w) = \frac{\mathcal{L}_f'(z)}{(\phi_f^\iota)'(z)} \cdot (C + \phi_f^\iota(z) - \phi_g^\iota(w)),$$

from which it follows directly that the index κ depends affinely on $\phi_g^\iota(w)$, although it is conceivable that the multiplicative constant in this dependence vanishes.

As will be explained in detail in Section 5C, the index κ is independent from w if and only if, denoting the fixed point of $\mathcal{L}_f$ again by z_0, we have

$$\frac{\mathcal{L}_f''(z_0)(\phi_f^\iota)'(z_0)}{\lambda(1-\lambda)} - (\phi_f^\iota)''(z_0) = 0, \tag{4}$$

in which case κ is constantly equal to $+1$. The second statement in Proposition B′ therefore implies that f can be chosen in order to obtain an inequality in (4), which implies that the affine dependence of κ on $\phi_g^\iota(w)$ is nonconstant.

It follows that there exists an open subset of $\mathcal{B}_g$ where the w-values are such that $\mathrm{Re}(\kappa)$ is strictly negative. Let $D_2 \subset \mathcal{B}_g$ be a small disk contained in this open subset, so that $\mathrm{Re}(\kappa)$ is negative for all $w \in D_2$.

Let D_1 be a small disk centered at z_0, the Siegel-type fixed point of $\mathcal{L}_f$. We claim that, for $n \in \mathbb{N}$ large enough, the open set $D_1 \times g^{n^2}(D_2)$ is contained in a wandering Fatou component.

Indeed, it follows from Proposition A′ that the nonautonomous one-dimensional system given by compositions of the maps $z \mapsto \pi_z \circ P^{2n+1}(z, g^{n^2}(w))$ satisfies case (2) of Theorem 2.2, where π_z is the projection onto the z-coordinate. Thus Theorem 2.2 implies that

$$P^{m^2-n^2}(z,w) \to (z_0, 0)$$

uniformly for all $(z,w) \in D_1 \times g^{n^2}(D_2)$. The remainder of the proof follows the argument from [Astorg et al. 2016]. Since the complement of the escape locus of P is bounded, it follows that the entire orbits $P^m(z,w)$ must remain uniformly bounded, which implies normality of (P^m) on $D_1 \times g^{n^2}(D_2)$, which is therefore contained in a Fatou component, say U. The fact that on an open subset of U the subsequence $P^{m^2-n^2}$ converges to the constant $(z_0,0)$ implies convergence of this subsequence to $(z_0,0)$ on all of U, since limit maps of convergent subsequences are holomorphic. But if U was periodic or preperiodic, the limit set would have been periodic. The point $(z_0,0)$ is however not periodic: its orbit converges to $(0,0)$. Thus U is wandering, which completes the proof.

Remark 2.4. From the above discussion we can conclude that all possible limit maps of the convergent subsequence $P^{n_j}|_U$ are points. In fact these points form (the closure of) a bi-infinite orbit of $(z_0,0)$, converging to $(0,0)$ both under backward and forward iteration.

We note however that there are fibers $\{w=w_0\}$, with $w_0\in\mathcal{B}_g$, for which $\mathrm{Re}(\kappa)=0$. Let D_1 again be a sufficiently small disk centered at z_0, the Siegel-type fixed point of $\mathcal{L}_f$. Proposition A$'$ together with case (1) of Theorem 2.2 implies that for sufficiently large n the disk $D_1\times\{g^{n^2}(w_0)\}$ is a Fatou disk for P, i.e., the restriction of the iterates P^n to the disk form a normal family. For this Fatou disk the sequence of iterates $P^{m^2-n^2}$ converges to a rank-1 limit map, whose image is a holomorphic disk containing $(z_0,0)$. All the limit sets together form (the closure of) a bi-infinite sequence of disks, converging under backward and forward iteration to the point $(0,0)$.

3. Perturbations of Siegel disks

3A. *Notation.* The following conventions will be used throughout this section:

(i) Given a holomorphic function f, we will write $\hat{f}$ for the nonlinear part of f.

(ii) For a sequence of constants $\lambda_n\in\mathbb{C}$ we will write

$$\lambda_{n,m}=\prod_{j=m+1}^{n}\lambda_j \quad\text{and}\quad \lambda(n)=\lambda_{n,0}=\prod_{j=1}^{n}\lambda_j,$$

and similarly for a sequence of functions (f_n)

$$f_{n,m}=f_n\circ\cdots\circ f_{m+1}.$$

(iii) Given two sequences of holomorphic functions (f_n) and (g_n) defined on some uniform neighborhood of the origin, we will write $f_n\asymp g_n$ if the norms of the sequence of differences (f_n-g_n) is summable on some uniform neighborhood of the origin.

3B. *Preparation.* In this section we introduce nonautonomous analogies of attracting, repelling, and locally linearizable indifferent fixed points and make a few initial observations. In the next subsection we introduce the index κ and show that the local behavior of the nonautonomous systems we consider can be deduced from the real part of the index.

Definition 3.1. Two sequences of functions (f_n) and (g_n) are said to be *nonautonomously conjugate* if there exist a uniformly bounded sequence of local coordinate changes $(\psi_n)_{n\geq n_0}$, all defined in a uniform neighborhood of the origin, satisfying

$$f_n\circ\psi_n=\psi_{n+1}\circ g_n$$

for all $n\geq n_0$.

Definition 3.2. A sequence of functions (f_n) is said to be nonautonomously *linearizable* if there exists a sequence $(\lambda_n)_{n\geq n_0}$ in $\mathbb{C}\setminus\{0\}$ and a sequence of coordinate changes $(\psi_n)_{n\geq n_0}$, defined and uniformly bounded in a uniform neighborhood of the origin, and with derivative $\psi_n'(0)$ uniformly bounded away from zero, so that

$$f_n \circ \psi_n(z) = \psi_{n+1}(\lambda_n \cdot z)$$

for all $n \geq n_0$. If the sequence $|\lambda(n)|$ is bounded, both from above and away from 0, then we say that (f_n) is *rotationally linearizable*.

Definition 3.3. A sequence of functions (f_n) is said to be *collapsing* if there is a neighborhood of the origin U and an $n_0 \in \mathbb{N}$ such that $f_{n,m} \to 0$ on U as $n \to \infty$ for any $m \geq n_0$.

An example of a collapsing sequence is given by a sequence of functions f_n converging to a function f with an attracting fixed point at the origin.

Definition 3.4. We say that sequence (f_n) is *expulsive* if there exists $r > 0$ such that for every $m \geq 0$ there exists at most one exceptional point $\hat{z}$ such that for every $z \in D_r(0)\setminus\{\hat{z}\}$ there exist $n > m$ for which $f_{n,m}(z) \notin D_r(0)$. Here $D_r(0)$ denotes the disk of radius r centered at the origin.

An example of an expulsive sequence can be obtained by considering a sequence of maps (f_n) converging locally uniformly to a map with a repelling fixed point. Since f maps a small disk around the origin to a strictly larger holomorphic disk, the same holds for sufficiently small perturbations. A nested sequence argument shows that, starting at a sufficiently large time n_0, there is a unique orbit which remains in the small disk.

Lemma 3.5. *Consider a sequence (f_n) of univalent holomorphic functions, defined in a uniform neighborhood of the origin. Suppose the compositions $f_{n,0}$ are all defined in a possibly smaller neighborhood of the origin, and form a normal family. Then the sequence (f_n) is either rotationally linearizable, or there exist subsequences (n_j) for which $f_{n_j,0}$ converges to a constant.*

Proof. By normality the orbit $f_{n,0}(0)$ stays bounded. By nonautonomously conjugating with a sequence of translations we may therefore assume that $f_n(0) = 0$ for all n. Note that normality is preserved under nonautonomous conjugation by bounded translations.

Write $\lambda_n = f_n'(0)$. Normality implies that $|\lambda(n)|$ is bounded from above. The functions

$$\psi_{n+1}(z) := f_{n,0}(\lambda(n)^{-1} \cdot z)$$

are tangent to the identity, and they satisfy the functional equation

$$f \circ \psi_n(z) = \psi_{n+1}(\lambda \cdot z).$$

If the sequence $|\lambda(n)|$ is bounded away from the origin then the maps ψ_n are uniformly bounded, and the sequence $(f(n))$ is rotationally linearizable. Suppose that the sequence $\lambda(n)$ is not bounded from below, in which case there is a subsequence $\lambda(n_j)$ converging to 0. By the Hurwitz theorem the sequence of maps $f_{n_j,0}$ converges to a constant. □

Lemma 3.6. *If the sequence (f_n) is rotationally linearizable, and (ζ_n) is a sequence of absolutely summable holomorphic functions, i.e.,*

$$\sum \|\zeta_n\|_{D_r(0)} < \infty$$

for some $r > 0$, then the sequence $(f_n + \zeta_n)$ is also rotationally linearizable.

Proof. Write $g_n = f_n + \zeta_n$. We consider the errors due to the perturbations in linearization coordinates, i.e.,

$$\psi_{n+1}^{-1} \circ g_n \circ \psi_n(z) - \psi_{n+1}^{-1} \circ f_n \circ \psi_n(z) = \psi_{n+1}^{-1} \circ g_n \circ \psi_n(z) - \lambda_n \cdot z.$$

By the definition of the nonautonomous linearization, it follows that after restricting to a smaller neighborhood of the origin the derivatives of the maps ψ_n and their inverses are uniformly bounded. It follows that the above errors are also absolutely summable, which guarantees normality of the sequence $\psi_{n+1}^{-1} \circ g_{n,0}$ in a small neighborhood of the origin, and hence normality of the sequence $g_{n,0}$. It follows from Lemma 3.5 that $(f_n + \zeta_n)$ is either rotationally linearizable or has subsequences converging to the origin. It follows from the summability of the errors that the latter is impossible. □

3C. *Introduction of the index.* Let $f(z) = \lambda z + b_2 z^2 + O(z^3)$ be a holomorphic function with $\lambda = e^{2\pi i \zeta}$ and $\zeta \in \mathbb{R} \backslash \mathbb{Q}$ Diophantine. Let $h(z) = c_0 + c_1 z + O(z^2)$ be a holomorphic function defined in a neighborhood of the origin. Let $(\zeta_n(z))$ be a sequence of holomorphic functions that is defined and absolutely summable on some uniform neighborhood of the origin. We consider the nonautonomous dynamical system given by compositions of the maps

$$f_n(z) = f(z) + \frac{1}{n} h(z) + \zeta_n(z).$$

We introduce the index κ, depending rationally on the two-jet of f at the origin and the one-jet of h at the origin, by

$$\kappa := \frac{2 b_2 c_0}{\lambda(1-\lambda)} + \frac{c_1}{\lambda}. \tag{5}$$

We claim that the index κ is invariant under local autonomous changes of coordinates, i.e., when all the maps f_n are conjugated by a single analytic transformation. One easily observes that the index is invariant under affine changes of coordinates and is unaffected by terms of order 3 and higher. It is therefore sufficient to only consider local changes of the form $z \mapsto z + \alpha z^2$. It is clear that λ and c_0 are unaffected by such a coordinate change, while computation shows that b_2 is replaced by $b_2 + \alpha\lambda - \alpha\lambda^2$ and c_1 is replaced by $c_1 - \alpha\lambda c_0$. Indeed, κ is invariant under these changes.

Since ζ is Diophantine, the function f is linearizable. Let us write $\phi(z) = z + \text{h.o.t.}$ for the linearization map of f, i.e., $f \circ \phi(z) = \phi(\lambda z)$.

We define

$$\theta_n(z) := z + \frac{1}{n} \frac{c_0}{1-\lambda}.$$

Lemma 3.7. *With the above definitions we can write*

$$\mathrm{f}_n := \phi^{-1} \circ \theta_{n+1}^{-1} \circ f_n \circ \theta_n \circ \phi = \lambda \cdot e^{\kappa/n} \cdot z + \frac{1}{n} \sum_{k=2}^{\infty} d_k z^k + \xi_n(z), \tag{6}$$

where (ξ_n) *is a sequence of holomorphic functions that are defined and whose norms are summable on a uniform neighborhood of the origin.*

Proof. First observe that

$$\begin{aligned}\theta_{n+1}^{-1}\circ f_n\circ\theta_n &\asymp f\left(z+\frac{1}{n}\frac{c_0}{1-\lambda}\right)+\frac{1}{n}h\left(z+\frac{1}{n}\frac{c_0}{1-\lambda}\right)-\frac{1}{n+1}\frac{c_0}{1-\lambda}\\ &\asymp f(z)+f'(z)\frac{1}{n}\frac{c_0}{1-\lambda}+\frac{1}{n}h(z)-\frac{1}{n+1}\frac{c_0}{1-\lambda}.\end{aligned}$$

Using the power series expansions of f' and h we can therefore write

$$\begin{aligned}\theta_{n+1}^{-1}\circ f_n\circ\theta_n &\asymp f(z)+\left(\frac{1}{n}\frac{\lambda c_0}{1-\lambda}+\frac{1}{n}c_0-\frac{1}{n+1}\frac{c_0}{1-\lambda}\right)+\frac{1}{n}\left(\frac{2b_2c_0}{1-\lambda}+c_1\right)z+\frac{1}{n}\sum_{k=2}^{\infty}\beta_k z^k\\ &\asymp f(z)+\frac{c_0}{1-\lambda}\left(\frac{1}{n}-\frac{1}{n+1}\right)+\frac{1}{n}\lambda\kappa z+\frac{1}{n}\sum_{k=2}^{\infty}\beta_k z^k\\ &\asymp f(z)+\frac{1}{n}\lambda\kappa z+\frac{1}{n}\sum_{k=2}^{\infty}\beta_k z^k.\end{aligned}$$

It follows that

$$\begin{aligned}\mathrm{f}_n &\asymp \phi^{-1}(f(\phi(z)))+(\phi^{-1})'(f(\phi(z)))\left(\frac{1}{n}\lambda\kappa\phi(z)+\frac{1}{n}\sum_{k=2}^{\infty}\beta_k\phi(z)^k\right)\\ &\asymp \lambda\left(1+\frac{\kappa}{n}\right)z+\frac{1}{n}\sum_{k=2}^{\infty}d_k z^k\\ &\asymp \lambda e^{\kappa/n}z+\frac{1}{n}\sum_{k=2}^{\infty}d_k z^k.\end{aligned}$$

For the last equality we used that

$$1+\frac{\kappa}{n}=e^{\kappa/n}+O\left(\frac{1}{n^2}\right).$$

□

Corollary 3.8. *If* $\mathrm{Re}(\kappa)<0$ *the sequence* f_n *is collapsing.*

Proof. Observe that $\mathrm{f}_n'(z)\asymp\lambda e^{\kappa/n}+O(z/n)$ and note that there is a small disk $D_r(0)$ such that for n sufficiently large

$$\|\mathrm{f}_n'\|_{D_r(0)}<e^{\mathrm{Re}(\kappa)/(2n)},$$

and thus

$$|\mathrm{f}_n(z)-\mathrm{f}_n(w)|\le e^{\mathrm{Re}(\kappa)/(2n)}|z-w|. \tag{7}$$

Since $\mathrm{Re}(\kappa)<0$ it follows that $\prod_{n\ge1}e^{\mathrm{Re}(\kappa)/(2n)}=0$.

Let us write

$$\varphi_n(z)=\lambda\cdot e^{\kappa/n}\cdot z+\hat{\mathrm{f}}_n(z),$$

i.e., we drop the term ξ_n from f_n. By decreasing the radius r if necessary we can choose m_0 such that

$$\sum_{j\ge m_0}\|\xi_j\|_{D_r(0)}<\tfrac{1}{2}r.$$

By increasing m_0 if necessary we can also guarantee that $\varphi_{n,m}(z) \in D_{r/2}(0)$ for all $z \in D_{r/4}(0)$ and $m \geq m_0$. Using (7) it follows by induction on n that whenever $z \in D_{r/4}(0)$ and $m \geq m_0$ we have

$$\|\mathrm{f}_{n,m}(z) - \varphi_{n,m}(z)\| \leq \sum_{j=m}^{n} \left(\prod_{k=j+1}^{n} e^{\mathrm{Re}(\kappa)/(2k)} \right) \|\xi_j\|_{D_r(0)}.$$

Indeed, the inequality is trivially satisfied for $n = m$, and assuming the inequality holds for some $n \geq m$ implies

$$\begin{aligned}\|\mathrm{f}_{n+1,m}(z) - \varphi_{n+1,m}(z)\| &= \|\mathrm{f}_{n+1} \circ \mathrm{f}_{n,m}(z) - \mathrm{f}_{n+1} \circ \varphi_{n,m}(z) + \xi_{n+1}(\varphi_{n,m}(z))\| \\ &\leq \sum_{j=m}^{n+1} \left(\prod_{k=j+1}^{n+1} e^{\mathrm{Re}(\kappa)/(2k)} \right) \|\xi_j\|_{D_r(0)}.\end{aligned}$$

Note that

$$\sum_{j=m}^{n} \left(\prod_{k=j+1}^{n} e^{\mathrm{Re}(\kappa)/(2k)} \right) \|\xi_j\|_{D_r(0)} \to 0$$

as $n \to \infty$; hence the fact that the sequence (φ_n) collapses implies the sequence (f_n) collapses as well. □

Since the sequence (f_n) collapses, it follows immediately that the sequence (f_n) collapses as well, concluding the case $\mathrm{Re}(\kappa) < 0$.

Corollary 3.9. *If* $\mathrm{Re}(\kappa) > 0$, *the sequence* f_n *is expulsive.*

Proof. Note that there are $r, n_0 > 0$ such that for every $z, w \in D_r(0)$ and every $n > n_0$ we have

$$|\mathrm{f}_n(z) - \mathrm{f}_n(w)| = |z - w| \cdot \left| e^{\kappa/n} + \frac{1}{n} O(z, w) \right| > e^{\mathrm{Re}(\kappa)/(2n)} |z - w|.$$

Expulsion of all but one orbit follows immediately. □

Again it follows that (f_n) is expulsive, completing the case $\mathrm{Re}(\kappa) > 0$.

3D. *Rotationally linearizable case* (Re(κ) = 0). Let us define

$$L_n(z) = e^{\kappa \log n} \cdot z.$$

We obtain

$$\begin{aligned} g_n &= L_{n+1}^{-1} \circ \mathrm{f}_n \circ L_n \\ &= \lambda z + e^{-\kappa \log(n+1)} \frac{1}{n} \sum_{\ell=2}^{\infty} d_\ell e^{\kappa \ell \log n} z^\ell + L_{n+1}^{-1} \circ \xi_n \circ L_n \\ &\asymp \lambda z + \frac{1}{n} \sum_{\ell=2}^{\infty} d_\ell e^{\kappa(\ell-1)\log n} z^\ell. \end{aligned}$$

Since $\mathrm{Re}(\kappa) = 0$, the maps L_n are rotations; hence it is sufficient to prove that the sequence (g_n) is rotationally linearizable.

By Lemma 3.6 we may ignore the absolutely summable part of g_n; hence with slight abuse of notation we may assume that

$$g_n = \lambda z + \frac{1}{n}\sum_{\ell=2}^{\infty} d_\ell e^{\kappa(\ell-1)\log n} z^\ell.$$

Recall that $\lambda = e^{2\pi i\zeta}$, where ζ is Diophantine.

Lemma 3.10. *There exist constants $C, r > 0$ such that for every integer $\ell \geq 1$ and for every $0 < m < N$ we have*

$$\left|\sum_{j=m}^{N} \lambda^{\ell j}\right| < C\ell^r.$$

Proof. Since ζ is assumed to be Diophantine, there exist $c, r > 0$ such that $|\lambda^n - 1| \geq cn^{-r}$ for all n. This gives the bound

$$\left|\sum_{j=m}^{N} \lambda^{\ell j}\right| = \left|\sum_{j=m}^{N} \frac{\lambda^{\ell(j+1)} - \lambda^{\ell j}}{\lambda^\ell - 1}\right| = \left|\frac{1}{\lambda^\ell - 1}\sum_{j=m}^{N}(\lambda^{\ell(j+1)} - \lambda^{\ell j})\right| < \left|\frac{2}{\lambda^\ell - 1}\right| < C\ell^r. \qquad \square$$

Lemma 3.11. *There exist $\widetilde{C}, r > 0$ such that for all integers $n, \ell > 0$ we have*

$$\left|\sum_{k=n}^{\infty} \frac{e^{\kappa\ell\log k}}{k}\lambda^{k\ell}\right| < \frac{\widetilde{C}\ell^{r+1}}{n}.$$

Proof. Summation by parts gives

$$\begin{aligned}
\sum_{k=n}^{N} \frac{e^{\kappa\ell\log k}}{k}\lambda^{k\ell} &= \frac{e^{\kappa\ell\log N}}{N}\sum_{k=n}^{N}\lambda^{k\ell} - \sum_{k=n}^{N-1}\left(\frac{e^{\kappa\ell\log(k+1)}}{k+1} - \frac{e^{\kappa\ell\log k}}{k}\right)\sum_{j=n}^{k}\lambda^{j\ell} \\
&= \frac{e^{\kappa\ell\log N}}{N}\sum_{k=n}^{N}\lambda^{k\ell} - \sum_{k=n}^{N-1} e^{\kappa\ell\log k}\left(\frac{1+\kappa l/k + O(1/k^2)}{k+1} - \frac{1}{k}\right)\sum_{j=n}^{k}\lambda^{j\ell}.
\end{aligned}$$

Observe that

$$\frac{1+\kappa\ell/k + O(1/k^2)}{k+1} - \frac{1}{k} = O\left(\frac{1}{k^2}\right)$$

is absolutely summable; hence using Lemma 3.10 we obtain

$$\begin{aligned}
\left|\sum_{k=n}^{N} \frac{e^{\kappa\ell\log k}}{k}\lambda^{k\ell}\right| &< \frac{1}{N}\left|\sum_{k=n}^{N}\lambda^{k\ell}\right| + \sum_{k=n}^{N-1}\left|\frac{1+\kappa\ell/k + O(1/k^2)}{k+1} - \frac{1}{k}\right|\left|\sum_{j=n}^{k}\lambda^{j\ell}\right| \\
&< \frac{C\ell^r}{N} + C\ell^r\sum_{k=n}^{N-1}\left|\frac{\ell\kappa - 1}{k(k+1)} + O\left(\frac{1}{k^3}\right)\right| \\
&< \frac{\widetilde{C}\ell^{r+1}}{n}. \qquad \square
\end{aligned}$$

Let us introduce one more change of coordinates

$$S_{n+1}(z) = z - \lambda^{-1} \sum_{\ell=2}^{\infty} \lambda^{(n+1)(1-\ell)} d_\ell z^\ell \sum_{k=n+1}^{\infty} \frac{e^{\kappa(\ell-1)\log k}}{k} \lambda^{k(\ell-1)}.$$

Lemma 3.12. *Writing $S_n(z) = z + \widehat{S}_n(z)$ we obtain*

$$\widehat{S}_{n+1}(\lambda z) = \lambda \widehat{S}_n(z) + \hat{g}_n(z).$$

Proof. Computing $\widehat{S}_{n+1}(\lambda z) - \hat{g}_n(z)$ gives

$$-\lambda^{-1} \sum_{\ell=2}^{\infty} \lambda^{(n+1)(1-\ell)} d_\ell \lambda^\ell z^\ell \sum_{k=n+1}^{\infty} \frac{e^{\kappa(\ell-1)\log k}}{k} \lambda^{k(\ell-1)} - \frac{1}{n} \sum_{\ell=2}^{\infty} e^{\kappa(\ell-1)\log n} d_\ell z^\ell$$

$$= -\sum_{\ell=2}^{\infty} \lambda^{n(1-\ell)} d_\ell z^\ell \sum_{k=n+1}^{\infty} \frac{e^{\kappa(\ell-1)\log k}}{k} \lambda^{k(\ell-1)} - \sum_{\ell=2}^{\infty} \frac{e^{\kappa(\ell-1)\log n}}{n} \lambda^{n(\ell-1)} \lambda^{n(1-\ell)} d_\ell z^\ell$$

$$= -\sum_{\ell=2}^{\infty} \lambda^{n(1-\ell)} d_\ell z^\ell \sum_{k=n}^{\infty} \frac{e^{\kappa(\ell-1)\log k}}{k} \lambda^{k(\ell-1)} = \lambda \widehat{S}_n(z). \qquad \square$$

Lemma 3.13. *The maps S_n satisfy $S_n = z + O(1/n)$, with uniform bounds.*

Proof.

$$|\widehat{S}_n(z)| = \left| \lambda^{-1} \sum_{\ell=2}^{\infty} \lambda^{n(1-\ell)} d_\ell z^\ell \sum_{k=n}^{\infty} \frac{e^{\kappa(\ell-1)\log k}}{k} \lambda^{k(\ell-1)} \right| < \frac{\widetilde{C}}{n} \sum_{\ell=2}^{\infty} |d_\ell z^\ell| (\ell-1)^{r+1}. \qquad \square$$

Let us define

$$h_n := S_{n+1}^{-1} \circ g_n \circ S_n.$$

Lemma 3.14. *The maps h_n are of the form*

$$h_n = \lambda z + O(n^{-2}).$$

Proof. The definition of h_n immediately gives that $h_n(z) = \lambda z + O(1/n)$,

$$g_n \circ S_n = S_{n+1} \circ h_n,$$

and thus

$$\lambda z + \lambda \widehat{S}_n(z) + \hat{g}_n(z + \widehat{S}_n) = \lambda z + \hat{h}_n(z) + \widehat{S}_{n+1}(\lambda z + \hat{h}_n),$$

which gives

$$\lambda \widehat{S}_n(z) + \hat{g}_n(z) + \hat{g}_n'(z)\widehat{S}_n(z) + O(\widehat{S}_n^2) = \hat{h}_n(z) + \widehat{S}_{n+1}(\lambda z) + \widehat{S}_{n+1}'(\lambda z)\hat{h}_n + O(\hat{h}_n^2).$$

Hence by Lemma 3.12 we obtain

$$\hat{g}_n'(z)\widehat{S}_n(z) + O(\widehat{S}_n^2) = \hat{h}_n(z)(1 + \widehat{S}_{n+1}'(\lambda z)) + O(\hat{h}_n^2).$$

Since $\hat{g}_n = O(1/n)$ and $\widehat{S}_n = O(1/n)$, we get

$$\hat{h}_n(z)(1 + \widehat{S}'_{n+1}(\lambda z)) + O(\hat{h}_n^2) = O\left(\frac{1}{n^2}\right).$$

Since $h_n(z) = \lambda z + O(1/n)$, it follows that $\hat{h}_n(z) = O(1/n^2)$. □

Lemma 3.6 implies that the sequence (h_n) is rotationally linearizable; hence the same holds for (g_n), (f_n) and finally (f_n), which completes the proof of Theorem 2.2.

Remark 3.15. The proof of Theorem 2.2 also works for more general perturbations, for example

$$f_n(z) \asymp f(z) + \frac{1}{n} h_1(z) + \frac{\log n}{n} h_2(z),$$

where h_1 and h_2 are holomorphic around the origin. In this case we have two indexes κ_j, $j \in \{1, 2\}$, that can be computed using (5), where constants c_0 and c_1 are the coefficients of the linear part of the Taylor series of h_j at the origin. The following is a general version of Theorem 2.2:

(1) If $\mathrm{Re}(\kappa_2) > 0$ then the sequence (f_n) is expulsive.

(2) If $\mathrm{Re}(\kappa_2) < 0$ then the sequence (f_n) is collapsing.

(3) If $\mathrm{Re}(\kappa_2) = 0$ and:

(a) $\mathrm{Re}(\kappa_1) > 0$, then the sequence (f_n) is expulsive.

(b) $\mathrm{Re}(\kappa_1) < 0$, then the sequence (f_n) is collapsing.

(c) $\mathrm{Re}(\kappa_1) = 0$, then the sequence (f_n) is rotationally linearizable; hence all limit maps have rank 1.

4. Existence of parabolic curves

The purpose of this section is to prove the following proposition.

Proposition 4.1. *Let $P(z, w) := \big(f(z) + \frac{\pi^2}{4} w, g(w)\big)$, with $f(z) = z + z^2 + bz^3 + O(z^4)$ and $g(w) = w - w^2 + O(w^3)$. Then P has at least three parabolic curves: one is contained in the invariant fiber $w = 0$ and is an attracting petal for f; the other two are graphs over the same petal $\mathcal{P}$ in the parabolic basin $\mathcal{B}_g$. Moreover they are of the form*

$$\zeta^{\pm}(w) = \pm c_1 \sqrt{w} + c_2 w \pm c_3 w^{3/2} + O(w^2),$$

where $c_1 = \frac{\pi}{2} i$ and $c_2 = \frac{\pi^2}{8} b - \frac{1}{4}$.

Proposition 4.1 gives a positive answer to a question posed in [Astorg et al. 2016]. We note that the result does not follow from the results [Hakim 1998], as the two characteristic directions we consider are *degenerate*, in the language used by Hakim. The existence of three parabolic curves can be derived from [López-Hernanz and Rosas 2020]. However, their proof gives no guarantee that the parabolic curves $\zeta^{\pm}$ are graphs over the same petal in $\mathcal{B}_g$, which is crucial for our purpose.

Let us start by observing that P is semiconjugate to a map Q, holomorphic near the origin, given by

$$Q(z, \epsilon) = \big(f(z) + \tfrac{\pi^2}{4} \epsilon^2, \epsilon - \tfrac{1}{2} \epsilon^3 + O(\epsilon^5)\big)$$

(with $\epsilon^2 = w$). The map Q has three characteristic directions: $\epsilon = 0$, $z = \frac{\pi}{2} i\epsilon$ and $z = -\frac{\pi}{2} i\epsilon$. It is clear that there is a parabolic curve tangent to the characteristic direction $\epsilon = 0$, namely the attracting petal for f in the invariant fiber $\{\epsilon = 0\}$. We call this parabolic curve the trivial curve. For the existence of the two other parabolic curves we will use a graph transform argument.

Let us write $Q(z, \epsilon) = (f_\epsilon(z), \tilde{g}(\epsilon))$, so that $f_\epsilon(z) = f(z) + \frac{\pi^2}{4}\epsilon^2$ and $\tilde{\epsilon} := \tilde{g}(\epsilon) = \sqrt{g(\epsilon^2)} = \epsilon - \frac{1}{2}\epsilon^3 + O(\epsilon^5)$. We are looking for parabolic curves of the form $\epsilon \to (\zeta(\epsilon), \epsilon)$, hence satisfying the equation

$$Q(\zeta(\epsilon), \epsilon) = (\zeta(\tilde{\epsilon}), \tilde{\epsilon}). \tag{8}$$

Equivalently we are looking for a function ζ, defined for ϵ in a parabolic petal of $\tilde{g}$, satisfying the functional equation

$$\zeta(\tilde{g}(\epsilon)) = f_\epsilon(\zeta(\epsilon)).$$

We will prove that Q has two parabolic curves $\zeta^\pm$, corresponding to the characteristic directions $z = \pm\frac{\pi}{2} i\epsilon$, which are graphs over the same attracting petal of $\tilde{g}$ in the right half-plane. This will complete the proof of Proposition 4.1, since these two parabolic curves can be lifted to parabolic curves of P satisfying the desired properties.

The key idea in proving the existence of $\zeta(\epsilon)$ is to start with sufficiently high-order jets $\zeta_1(\epsilon)$ of the *formal* solution to (8), and then apply a graph transform argument, starting with ζ_1. By starting with higher-order jets, we obtain higher-order error estimates, but the constants in those estimates are likely to deteriorate. However, these estimates can be controlled by dropping the order of the error estimates by 1, and working with $|\epsilon| < \delta$, with δ depending on the order of the jets. It turns out that starting with jets of order 20 is sufficient to obtain convergence of the graph transforms. We do not claim that 20 is the minimal order for which convergence can be obtained, only that the order suffices for our purposes.

Lemma 4.2. *For every integer $n > 0$ there exists $\zeta_1(\epsilon) = c_1\epsilon + c_2\epsilon^2 + c_3\epsilon^2 + \cdots + c_n\epsilon^n$ and $\delta > 0$ such that $|\zeta_1(\tilde{\epsilon}) - f_\epsilon(\zeta_1(\epsilon))| < |\epsilon|^n$ for all $|\epsilon| < \delta$. Moreover we have $c_1 = \pm\frac{\pi}{2} i$ and $c_2 = \frac{\pi^2}{8} b - \frac{1}{4}$.*

Proof. Recall from [Astorg et al. 2016] that by choosing $\zeta_1(\epsilon) = c_1\epsilon + c_2\epsilon^2$, with $c_1 = \pm\frac{\pi}{2} i$ and $c_2 = \frac{\pi^2}{8} b - \frac{1}{4}$, we obtain

$$|\zeta_1(\tilde{\epsilon}) - f_\epsilon(\zeta_1(\epsilon))| < O(|\epsilon|^4).$$

Now suppose that $c_1, \ldots, c_n$ are found such that for $\zeta(\epsilon) = c_1\epsilon + \cdots + c_n\epsilon^n$ we have

$$|\zeta_1(\tilde{\epsilon}) - f_\epsilon(\zeta_1(\epsilon))| < O(|\epsilon|^{n+2}).$$

Let $E_n(\epsilon) := f_\epsilon(\zeta_1(\epsilon)) - \zeta_1(\tilde{\epsilon})$. For $c_{n+1} \in \mathbb{C}$, let

$$E_{n+1}(\epsilon) := f_\epsilon(\zeta_1(\epsilon) + c_{n+1}\epsilon^{n+1}) - \zeta_1(\tilde{\epsilon}) - c_{n+1}\tilde{\epsilon}^{n+1};$$

we shall prove that there exists some c_{n+1} such that $E_{n+1} = O(\epsilon^{n+3})$. Indeed,

$$\begin{aligned} f_\epsilon(\zeta_1(\epsilon) + c_{n+1}\epsilon^{n+1}) &= f_\epsilon(\zeta_1(\epsilon)) + f'(\zeta_n(\epsilon))c_n\epsilon^{n+1} + O(\epsilon^{2n+2}) \\ &= f_\epsilon(\zeta_1(\epsilon)) + (1 + 2c_1\epsilon)c_{n+1}\epsilon^{n+1} + O(\epsilon^{n+3}). \end{aligned}$$

On the other hand, we have $c_{n+1}\tilde{\epsilon}^{n+1} = c_{n+1}\epsilon^{n+1} + O(\epsilon^{n+3})$; so

$$E_{n+1}(\epsilon) = E_n(\epsilon) + 2c_1c_{n+1}\epsilon^{n+2} + O(\epsilon^{n+3}).$$

Since $E_n(\epsilon) = O(\epsilon^{n+2})$ (and $c_1 \neq 0$), we may therefore find some value of c_{n+1} for which $E_{n+1}(\epsilon) = O(\epsilon^{n+3})$.

We conclude that if δ is small enough then $|\zeta_1(\tilde{\epsilon}) - f_\epsilon(\zeta_1(\epsilon))| < |\epsilon|^n$ for all $|\epsilon| < \delta$. □

Remark 4.3. The choice of parabolic curve is determined by the choice of c_1. From now on we will assume that $c_1 = \frac{\pi}{2}i$; for the case $c_1 = -\frac{\pi}{2}i$ the proofs are essentially the same.

For $R \in \mathbb{C}$ we write $\mathbb{H}_R = \left\{Z \in \mathbb{C} : \arg(Z - R) \in \left(-\frac{\pi}{2} - \epsilon_0, \frac{\pi}{2} + \epsilon_0\right)\right\}$ for some $\epsilon_0 > 0$, and

$$\mathcal{P}_\delta = \left\{\epsilon \in \mathbb{C} : \epsilon^{-2} \in \mathbb{H}_{\delta-2} \text{ and } \operatorname{Re}(\epsilon) > 0\right\}.$$

For $\delta > 0$ sufficiently small the petal $\mathcal{P}_\delta$ is forward-invariant under $\tilde{g}$, i.e., $\tilde{g}(\mathcal{P}_\delta) \subset \mathcal{P}_\delta$. Recall the existence of Fatou coordinates on P_δ: the function $\tilde{g}$ is conjugate to the translation $T_1 : Z \mapsto Z + 1$ via a conjugation of the form

$$Z = \frac{1}{\epsilon^2} + \alpha \log(\epsilon) + o(1),$$

where the constant α depends on g. All forward orbits in P_δ converge to 0 tangent to the positive real axis, and the conjugation gives the estimates

$$|\operatorname{Re}(\tilde{g}^k(\epsilon))| < \frac{C}{\sqrt{k}} \quad \text{and} \quad |\operatorname{Im}(\tilde{g}^k(\epsilon))| < \frac{C}{k} \tag{9}$$

for a uniform $C > 0$ depending on δ. We note that by choosing δ sufficiently small, the constant C can be chosen arbitrarily small as well.

Lemma 4.4. *Let $n > 0$ and $\zeta_1(\epsilon)$ be as in Lemma 4.2. There exist $\delta, A > 0$ such that for every $|\epsilon| < \delta$ we have*

$$|f^{-1}(f(\zeta_1(\epsilon)) + 3\epsilon^4) - \zeta_1(\epsilon)| \leq A|\epsilon|^4.$$

Proof. The Taylor series expansion of f gives

$$|f^{-1}(f(\zeta_1(\epsilon)) + 3\epsilon^4) - \zeta_1(\epsilon)| \leq \sum_{i=1}^{\infty} \left| \frac{(f^{-1})^{(i)}(f(\zeta_1(\epsilon)))}{i!} \right| 3^i |\epsilon^4|^i,$$

and the desired estimate follows immediately. □

Lemma 4.5. *Let $n > 0$ and $\zeta_1(\epsilon)$ be as in Lemma 4.2, $A > 0$ and $\delta > 0$ sufficiently small. Let $(\zeta_k(\epsilon))$ be any sequence of holomorphic functions defined on $\mathcal{P}_\delta$ and satisfying*

$$|\zeta_k(\epsilon) - \zeta_1(\epsilon)| < A|\epsilon|^4.$$

Then there exists $C_1 > 0$, depending on ζ_1, such that

$$\left| \prod_{s=\ell}^{k} f'(\zeta_s(\tilde{g}^{k+1-s}(\epsilon))) \right|^{-1} < C_1 \cdot (k + 1 - \ell)$$

for all $\epsilon \in \mathcal{P}_\delta$ and every $0 < \ell \leq k$.

Proof. Let us write $x_k = \mathrm{Re}(\tilde{g}^k(\epsilon)) > 0$ and $y_k = \mathrm{Im}(\tilde{g}^k(\epsilon))$. Estimates (9) imply

$$\sum_{k=0}^{\infty} |\tilde{g}^k(\epsilon)|^3 < K < \infty \quad \text{for all } \epsilon \in \mathcal{P}_\delta. \tag{10}$$

Since by assumption $|\zeta_s(\epsilon) - \zeta_1(\epsilon)| < A|\epsilon|^4$ for every $s \geq 1$, it follows that $\zeta_s(\epsilon) = c_1\epsilon + c_2\epsilon^2 + O(\epsilon^3)$ and

$$|f'(\zeta_s(\epsilon)) - f'(\zeta_1(\epsilon))| < B|\epsilon^4|,$$

where $B > 0$ depends only on ζ_1 and A.

Observe that $f'(z) = 1 + 2z + 3bz^2 + O(z^3) = e^{2z+(3b-2)z^2+O(z^3)}$; hence we obtain

$$f'(\zeta_s(\epsilon)) = e^{\pi i \epsilon + (\frac{1}{2}\pi^2(1-b) - \frac{1}{2})\epsilon^2 + O(\epsilon^3)},$$

where the bound $O(\epsilon^3)$ is uniform with respect to s.

Therefore we can find $C_1 > 0$ such that

$$\begin{aligned}
\left|\prod_{s=\ell}^{k} f'(\zeta_s(\tilde{g}^{k+1-s}(\epsilon)))\right| &> \left|e^{\sum_{s=\ell}^{k} \mathrm{Re}\left(\pi i \tilde{g}^{k+1-s}(\epsilon) + (\frac{1}{2}\pi^2(1-b)-\frac{1}{2})(\tilde{g}^{k+1-s}(\epsilon))^2\right) + O((\tilde{g}^{k+1-s}(\epsilon))^3)}\right| \\
&> \frac{1}{C_1}\left|e^{\sum_{s=\ell}^{k} -\pi y_{k+1-s} + (\frac{1}{2}\pi^2(1-\mathrm{Re}(b))-\frac{1}{2})x_{k+1-s}^2}\right| \\
&> \frac{1}{C_1}\left|e^{-\sum_{s=\ell}^{k} 1/(k+1-s)}\right| \\
&> \frac{1}{C_1(k+1-\ell)}.
\end{aligned}$$

In the first inequality we used the fact that $|e^z| = e^{\mathrm{Re}(z)}$. The second inequality follows from estimates (9) and (10). The third inequality depends on the constant C from (9) being sufficiently small, which can be guaranteed by taking sufficiently small δ. □

Remark 4.6. Note that the estimates in Lemmas 4.2, 4.4 and 4.5 hold regardless of the choice of n in the definition of ζ_1. If n is increased, then all estimates hold, with the same constants, for δ sufficiently small. It turns out that it will be sufficient for us to work with $n = 20$, and we will work with this choice from now on.

Lemma 4.7. *There exists sufficiently small $\delta > 0$ such that for every $k \geq 2$ and every $\epsilon \in \mathcal{P}_\delta$ we have*

$$|\tilde{g}^k(\epsilon)|^{19}k + |\tilde{g}^k(\epsilon)|^{39}(k-1) + \sum_{\ell=2}^{k-1} \frac{|\tilde{g}^{k+1-\ell}(\epsilon)|^{23}(k-\ell)}{(\ell-1)^4} + \frac{|\tilde{g}(\epsilon)|^{23}}{(k-1)^4} < \frac{4|\epsilon|^{12}}{k^2}.$$

Proof. We will prove that each of the four terms in the left-hand summation is bounded by $|\epsilon|^{12}/k^2$. It follows from (9) that for every $0 \leq \ell \leq 19$ we have

$$|\tilde{g}^k(\epsilon)|^{19}k < \frac{C^{19}k}{|k + 1/\epsilon^2|^{19/2}} < \frac{C^{19}|\epsilon|^\ell k}{k^{(19-\ell)/2}}.$$

If we choose $\ell = 13$ and assume that δ is small enough, then we get

$$|\tilde{g}^k(\epsilon)|^{19}k < \frac{|\epsilon|^{12}}{k^2}$$

for $\epsilon \in \mathcal{P}_\delta$. The desired bound for a second term follows immediately from the inequality

$$|\tilde{g}^k(\epsilon)|^{39}(k-1) < |\tilde{g}^k(\epsilon)|^{19}k.$$

Next observe that for every $k \geq 2$ we have

$$\frac{|\tilde{g}(\epsilon)|^{23}}{(k-1)^4} < \frac{2^2|\epsilon|^{23}}{(2(k-1)^2)^2} < \frac{|\epsilon|^{12}}{k^2},$$

where the last inequality holds for sufficiently small δ. Finally, for the third term in the summation we use (9) to obtain

$$\sum_{\ell=2}^{k-1} \frac{|\tilde{g}^{k+1-\ell}(\epsilon)|^{23}(k-\ell)}{(\ell-1)^4} < \sum_{\ell=2}^{k-1} \frac{C^{10}|\epsilon|^{13}(k-\ell)}{(k+1-\ell)^5(\ell-1)^4} < C^{10}|\epsilon| \sum_{\ell=2}^{k-1} \frac{|\epsilon|^{12}}{(k-\ell)^4(\ell-1)^4}.$$

In order to obtain the desired bound it suffices to prove that

$$\sum_{\ell=2}^{k-1} \frac{1}{(k-\ell)^4(\ell-1)^4} < \frac{4^4}{k^2}.$$

First observe that

$$\frac{1}{(\ell-1)(k-\ell)} \leq \frac{4}{k}$$

for every $k \geq 3$ and $2 \leq \ell \leq k-1$. To see this let us set $s = \ell - 1$ and $t = k-1$. The above inequality now translates to

$$\frac{1}{s(t-s)} \leq \frac{4}{t+1}$$

for $t \geq 2$ and $1 \leq s \leq t-1$, and hence to

$$p_t(s) := 4s^2 - 4ts + t + 1 \leq 0.$$

Observe that $p_t(1) < 0$ and that roots of $p_t(s)$ lie outside the closed interval $[1, t-1]$. Therefore we obtain

$$\sum_{\ell=2}^{k-1} \frac{1}{(k-l)^4(\ell-1)^4} < \sum_{\ell=2}^{k-1} \frac{4^4}{k^4} < \frac{4^4}{k^2},$$

and hence for δ sufficiently small

$$\sum_{\ell=2}^{k-1} \frac{|\tilde{g}^{k+1-\ell}(\epsilon)|^{23}(k-\ell)}{(\ell-1)^4} < \frac{|\epsilon|^{12}}{k^2}.$$

□

Proof of Proposition 4.1. As we remarked at the beginning of this section, it is enough to prove that Q has two parabolic curves $\zeta^\pm$ corresponding to the characteristic directions $z = \pm\frac{\pi}{2}i\epsilon$, both curves graph over the same attracting petal of $\tilde{g}$ in the right half-plane. By Lemma 4.2 there exist $\delta > 0$ and $\zeta_1(\epsilon) = c_1\epsilon + c_2\epsilon^2 + \cdots + c_{20}\epsilon^{20}$ such that $|\zeta_1(\tilde{\epsilon}) - f_\epsilon(\zeta_1(\epsilon))| < |\epsilon|^{20}$ for $|\epsilon| < \delta$. Let $A > 0$ be as in Lemma 4.4, and let $C_1 > 0$ be the constant defined in Lemma 4.5.

We will show that the sequence of functions defined inductively by

$$\zeta_{k+1}(\epsilon) := f_\epsilon^{-1}(\zeta_k(\tilde{\epsilon}))$$

is convergent and that the limit satisfies the functional equation (8). Let us define

$$E_k(\epsilon) := \zeta_k(\tilde{\epsilon}) - f_\epsilon(\zeta_k(\epsilon))$$

and observe that

$$\zeta_{k+1}(\epsilon) = f_\epsilon^{-1}(\zeta_k(\tilde{\epsilon})) = f_\epsilon^{-1}(f_\epsilon(\zeta_k(\epsilon)) + E_k(\epsilon))$$

and hence

$$f_\epsilon(\zeta_{k+1}(\epsilon)) = f_\epsilon(\zeta_1(\epsilon)) + \sum_{\ell=1}^{k} E_\ell(\epsilon).$$

Note that we can replace f_ϵ by f on both sides, giving

$$\zeta_{k+1}(\epsilon) = f^{-1}\left(f(\zeta_1(\epsilon)) + \sum_{\ell=1}^{k} E_\ell(\epsilon)\right),$$

and hence

$$\zeta_{k+1}(\epsilon) = \zeta_1(\epsilon) + \sum_{i=1}^{\infty} \frac{(f^{-1})^{(i)}(f(\zeta_1(\epsilon)))}{i!}\left(\sum_{\ell=1}^{k} E_\ell(\epsilon)\right)^i.$$

We will prove that $|E_k(\epsilon)| < |\epsilon|^4/|k-1|^2$ for every $k \geq 2$ on some small petal $\mathcal{P}_\delta$. This will imply that the sequence ζ_{k+1} converges to a parabolic curve ζ on $\mathcal{P}_\delta$ for sufficiently small δ.

We claim there exists $\delta > 0$ such that for every $\epsilon \in \mathcal{P}_\delta$ and every $k > 1$ the following two statements hold:

$I_k(1)$: $|\zeta_k(\epsilon) - \zeta_1(\epsilon)| < A|\epsilon|^4$, and

$I_k(2)$: $|E_k(\epsilon)| < 4|\epsilon|^{12}/|k-1|^2 < |\epsilon|^4/|k-1|^2$.

We will prove these two statements simultaneously by induction on k.

Step 1: First we prove $I_2(1)$. By definition

$$\zeta_2 = \zeta_1(\epsilon) + \sum_{i=1}^{\infty} \frac{(f^{-1})^{(i)}(f(\zeta_1(\epsilon)))}{i!}(E_1(\epsilon))^i;$$

hence by Lemma 4.4 we obtain the desired inequality.

Next we prove that $I_2(2)$. Observe that for sufficiently small δ we get

$$\begin{aligned}|E_2(\epsilon)| &< \left|\frac{E_1(\tilde{\epsilon})}{f'(\zeta_1(\tilde{\epsilon}))}\right| + C_2|E_1(\tilde{\epsilon})|^2 < C_1|\tilde{g}(\epsilon)|^{20} + C_2|\epsilon|\,|\tilde{g}(\epsilon)|^{40} \\ &< C_1|\epsilon|\,|\tilde{g}(\epsilon)|^{19} + C_2|\epsilon|\,|\tilde{g}(\epsilon)|^{39} < 4|\epsilon|^{12}.\end{aligned}$$

Here C_1 is the constant introduced in Lemma 4.5.

Step 2: Now let us assume that $I_\ell(1)$ and $I_\ell(2)$ hold for every $2 \le \ell \le k$. Observe that

$$|\zeta_{k+1}(\epsilon)-\zeta_1(\epsilon)| < \sum_{i=1}^{\infty}\left|\frac{(f^{-1})^{(i)}(f(\zeta_1(\epsilon)))}{i!}\right|\left|\sum_{\ell=1}^{k}E_\ell(\epsilon)\right|^i.$$

Since $|E_\ell(\epsilon)| < |\epsilon|^4/|\ell-1|^2$ for $\ell \ge 2$ and $|E_1(\epsilon)| < |\epsilon|^4$, we get

$$\left|\sum_{\ell=1}^{k}E_\ell(\epsilon)\right| < 3|\epsilon|^4;$$

hence by Lemma 4.4 inequality $I_{k+1}(1)$ holds.

Observe that

$$\begin{aligned}E_{k+1}(\epsilon) &= \zeta_{k+1}(\tilde\epsilon) - f_\epsilon(\zeta_{k+1}(\epsilon)) = \zeta_{k+1}(\tilde\epsilon)-\zeta_k(\tilde\epsilon)\\ &= f_{\tilde\epsilon}^{-1}(f_{\tilde\epsilon}(\zeta_k(\tilde\epsilon)) + E_k(\tilde\epsilon)) - \zeta_k(\tilde\epsilon)\\ &= f^{-1}(f(\zeta_k(\tilde\epsilon)) + E_k(\tilde\epsilon)) - \zeta_k(\tilde\epsilon)\\ &= (f^{-1})'(f(\zeta_k(\tilde\epsilon)))\cdot E_k(\tilde\epsilon) + O(E_k(\tilde\epsilon)^2),\end{aligned}$$

where the constant in the order can be chosen independently from k. It follows that there exists $C_2 > 0$ independent of k such that

$$|E_{k+1}(\epsilon)| < \left|\frac{E_k(\tilde\epsilon)}{f'(\zeta_k(\tilde\epsilon))}\right| + C_2|E_k(\tilde\epsilon)|^2. \tag{11}$$

Using the inequality (11) successively we obtain

$$\begin{aligned}|E_{k+1}(\epsilon)| &< \left|\frac{E_k(\tilde g(\epsilon))}{f'(\zeta_k(\tilde g(\epsilon)))}\right| + C_2|E_k(\tilde g(\epsilon))|^2\\ &< \left|\frac{E_{k-1}(\tilde g^2(\epsilon))}{f'(\zeta_{k-1}(\tilde g^2(\epsilon)))\cdot f'(\zeta_k(\tilde g(\epsilon)))}\right| + C_2\frac{|E_{k-1}(\tilde g^2(\epsilon))|^2}{|f'(\zeta_k(\tilde g(\epsilon)))|} + C_2|E_k(\tilde g(\epsilon))|^2\\ &< \frac{|E_1(\tilde g^k(\epsilon))|}{\prod_{\ell=1}^k |f'(\zeta_\ell(\tilde g^{k+1-\ell}(\epsilon)))|} + C_2\sum_{\ell=1}^{k-1}\frac{|E_\ell(\tilde g^{k+1-\ell}(\epsilon))|^2}{\prod_{s=\ell+1}^k |f'(\zeta_s(\tilde g^{k+1-s}(\epsilon)))|} + C_2|E_k(\tilde\epsilon)|^2.\end{aligned} \tag{12}$$

Combining (12) and Lemma 4.5 gives

$$\begin{aligned}&|E_{k+1}(\epsilon)|\\ &< C_1|\tilde g^k(\epsilon)|^{20}k + C_1C_2|\tilde g^k(\epsilon)|^{40}(k-1) + 16C_1C_2\sum_{\ell=2}^{k-1}\frac{|\tilde g^{k+1-\ell}(\epsilon)|^{24}(k-\ell)}{(\ell-1)^4} + 16C_1\frac{|\tilde g(\epsilon)|^{24}}{(k-1)^4}\\ &< C_1|\epsilon||\tilde g^k(\epsilon)|^{19}k + C_1C_2|\epsilon||\tilde g^k(\epsilon)|^{39}(k-1) + 16C_1C_2|\epsilon|\sum_{\ell=2}^{k-1}\frac{|\tilde g^{k+1-\ell}(\epsilon)|^{23}(k-\ell)}{(\ell-1)^4} + 16C_1|\epsilon|\frac{|\tilde g(\epsilon)|^{23}}{(k-1)^4}.\end{aligned}$$

If δ is sufficiently small this last inequality together with Lemma 4.7 implies

$$|E_{k+1}(\epsilon)| < \frac{4|\epsilon|^{12}}{k^2} < \frac{|\epsilon|^4}{k^2},$$

completing the proof of $I_{k+1}(2)$ and thus the induction argument. We emphasize that throughout the proof δ can be chosen dependently of k.

To summarize, the equation

$$\zeta_{k+1}(\epsilon) = f^{-1}\Bigl(f(\zeta_1(\epsilon)) + \sum_{\ell=1}^{k} E_\ell(\epsilon)\Bigr)$$

implies that for sufficiently small δ the sequence ζ_k converges on $\mathcal{P}_\delta$ to a parabolic curve ζ satisfying $\zeta(\tilde{\epsilon}) = f_\epsilon(\zeta(\epsilon))$. Recall that we have only proven the existence of a parabolic curve for $c_1 = \frac{\pi}{2}i$. For $c_1 = -\frac{\pi}{2}i$ we can use same arguments as above, but we might get a different value for δ. Since the parabolic petals are nested and forward invariant, both parabolic curves are graphs over the petal with minimal δ. $\square$

From the proof it follows that

$$\zeta^{\pm}(\epsilon) = \pm c_1\epsilon + c_2\epsilon^2 \pm c_3\epsilon^3 + O(\epsilon^4),$$

where $c_1 = \frac{\pi}{2}i$ and $c_2 = \frac{\pi^2}{8}b - \frac{1}{4}$.

5. Estimates on convergence towards Lavaurs map

5A. *Preliminaries.* The goal of this section is to obtain explicit estimates for one of the main objects to appear in our arguments: the functions $A(\epsilon, z)$ and $A_0(z)$, which measure how much the dynamics differ from a translation after a certain change of coordinates. The key difference between this section and the corresponding computations in [Astorg et al. 2016] is that we now know that we have two exactly invariant parabolic curves $\zeta^{\pm}$, instead of invariant jets. This is used crucially in the proof of Proposition 5.5.

Definition 5.1. Let $f_w(z) := f(z) + \frac{\pi^2}{4}w$, where $f(z) = z + z^2 + z^3 + O(z^4)$ is a degree-d polynomial. Let $g(w) = w - w^2 + O(w^3)$ be a degree-d polynomial.

In what follows, we set $\epsilon := \sqrt{w}$, working throughout with the branch that takes positive values on the positive real axis. We note that this branch is well-defined on the parabolic basin of the polynomial g. Abusing notation, we write $f_\epsilon(z) := f(z) + \frac{\pi^2}{4}\epsilon^2$ and $\zeta^{\pm}(\epsilon) = \pm i\frac{\pi}{2}\epsilon + c_2\epsilon^2 + O(\epsilon^3)$, where $\zeta^{\pm}$ are the parabolic curves constructed in the preceding section. Let $\tilde{g}(\epsilon) := \sqrt{g(\epsilon^2)} = \epsilon - \frac{1}{2}\epsilon^3 + O(\epsilon^5)$ ($\tilde{g}$ is analytic near $\epsilon = 0$).

Let us first record here the following lemma for later use:

Lemma 5.2. *Let $w_0 \in \mathcal{B}_g$ and let $\epsilon_j := \sqrt{g^{n^2+j}(w_0)}$. For $1 \le j \le n$, we have*

$$\epsilon_j = \frac{1}{n} - \frac{j}{2n^3} - \frac{\phi_g^\iota(w_0)}{2n^3} + o\Bigl(\frac{1}{n^3}\Bigr).$$

Proof. Let us write $w_{n^2+j} := g^{n^2+j}(w_0)$. We have

$$\phi_g(w_{n^2+j}) = \phi_g^\iota(w_0) + n^2 + j = \frac{1}{w_{n^2+j}} + o(1)$$

(note that we assume here $g(w) = w - w^2 + w^3 + O(w^4)$). Therefore

$$w_{n^2+j} = \frac{1}{n^2 + j + \phi_g^\iota(w_0) + o(1)},$$

and

$$\begin{aligned}\epsilon_j = \sqrt{w_{n^2+j}} &= \frac{1}{n}\left(1 + \frac{j + \phi_g^\iota(w_0)}{n^2} + o\left(\frac{1}{n^2}\right)\right)^{-1/2} \\ &= \frac{1}{n}\left(1 - \frac{1}{2}\frac{j + \phi_g^\iota(w_0)}{n^2} + o\left(\frac{1}{n^2}\right)\right).\end{aligned}$$ □

Definition 5.3. Let

$$\psi_\epsilon^\iota(z) := \frac{1}{i\pi}\log\left[\frac{\zeta^+(\epsilon) - z}{z - \zeta^-(\epsilon)}\right] + 1,$$
$$\psi_\epsilon^o(z) := \frac{1}{i\pi}\log\left[\frac{\zeta^+(\epsilon) - z}{z - \zeta^-(\epsilon)}\right] - 1,$$

where log is the principal branch of the logarithm.

Note with that choice of branch, ψ is defined on $\mathbb{C}\backslash L_\epsilon$, where L_ϵ is the real line through $\zeta^+(\epsilon)$ and $\zeta^-(\epsilon)$ minus the segment $[\zeta^-(\epsilon), \zeta^+(\epsilon)]$. In particular, ψ_ϵ^ι and ψ_ϵ^o are both defined in a disk centered at $z = 0$ whose radius is of order ϵ.

It will also be useful to note that

$$(\psi_\epsilon^{\iota/o})^{-1}(Z) = \frac{\zeta^+(\epsilon) - e^{\pm i\pi Z}\zeta^-(\epsilon)}{1 - e^{\pm i\pi Z}} = -\frac{\pi}{2}\epsilon\cot\left(\pm\frac{\pi Z}{2}\right) + O(\epsilon^2). \tag{13}$$

Definition 5.4. Let

(1) $A(\epsilon, z) := \psi_{\tilde{g}(\epsilon)}^{\iota/o} \circ f_\epsilon(z) - \psi_\epsilon^{\iota/o}(z) - \epsilon$,
(2) $A_0(z) := -1/f(z) + 1/z - 1$.

Note that the formula for $A(\epsilon, z)$ does not depend on whether the ingoing or outgoing coordinate ψ_ϵ is used, and is therefore well-defined.

Proposition 5.5. *We have*

(1) *A_0 is analytic near zero,*
(2) *there exists $r > 0$ such that for all $\epsilon \neq 0$ in a neighborhood of zero, $A(\epsilon, \cdot)$ is analytic on $\mathbb{D}(0, r)$.*

Proof. (1) A quick computation shows that

$$A_0(z) = \frac{f(z) - z - zf(z)}{zf(z)} = \frac{O(z^2)}{1 + O(z)},$$

from which the conclusion easily follows.

For (2), note that

$$
\begin{aligned}
A(\epsilon, z) &= \frac{1}{i\pi}\log\left(\frac{\zeta^+(\tilde{g}(\epsilon)) - f_\epsilon(z)}{f_\epsilon(z) - \zeta^-(\tilde{g}(\epsilon))} : \frac{\zeta^+(\epsilon) - z}{z - \zeta^-(\epsilon)}\right) - \epsilon \\
&= \frac{1}{i\pi}\log\left(\frac{f_\epsilon(\zeta^+(\epsilon)) - f_\epsilon(z)}{\zeta^+(\epsilon) - z} : \frac{f_\epsilon(z) - f_\epsilon(\zeta^-(\epsilon))}{z - \zeta^-(\epsilon)}\right) - \epsilon.
\end{aligned}
$$

From the above expression we see that the singularities at $z = \zeta^\pm(\epsilon)$ are in fact removable, unless one of the points coincides with a critical point of f. The fact that these critical points are bounded away from zero completes the proof. □

Lemma 5.6. *Let K be a compact subset of $\mathbb{C}^*$. There exists $C = C_K > 0$ such that, for all $z \in K$,*

$$
\left|\frac{z - \zeta^+(\epsilon)}{z - \zeta^-(\epsilon)} - \left(1 - \frac{i\pi}{z}\epsilon - \frac{\pi^2}{2z^2}\epsilon^2\right)\right| \leq C\epsilon^3.
$$

Proof. For $z \in K$, we have

$$
\begin{aligned}
\frac{\zeta^+(\epsilon) - z}{z - \zeta^-(\epsilon)} &= \frac{\zeta^+(\epsilon)}{z - \zeta^-(\epsilon)} - \frac{z}{z - \zeta^-(\epsilon)} \\
&= \frac{\zeta^+(\epsilon)}{z}\left(\frac{1}{1 - \zeta^-(\epsilon)/z}\right) - \frac{1}{1 - \zeta^-(\epsilon)/z} \\
&= \frac{\zeta^+(\epsilon)}{z}\left(1 + \frac{\zeta^-(\epsilon)}{z} + O(\epsilon^2)\right) - \left(1 + \frac{\zeta^-(\epsilon)}{z} + \left(\frac{\zeta^-(\epsilon)}{z}\right)^2 + O(\epsilon^3)\right) \\
&= \frac{c_1\epsilon + c_2\epsilon^2 - \frac{c_1^2}{z}\epsilon^2}{z} - 1 - \frac{-c_1\epsilon + c_2\epsilon^2}{z} - \frac{c_1^2\epsilon^2}{z^2} + O(\epsilon^3) \\
&= -1 + \frac{2c_1}{z}\epsilon - \frac{2c_1^2}{z^2}\epsilon^2 + O(\epsilon^3) \\
&= -1 + \frac{i\pi}{z}\epsilon + \frac{\pi^2}{2z^2}\epsilon^2 + O(\epsilon^3).
\end{aligned}
$$

□

Lemma 5.7. *Let K be a compact subset of $\mathbb{C}^*$. Then*

$$
\frac{f_\epsilon(z) - f_\epsilon(\zeta^+(\epsilon))}{f_\epsilon(z) - f_\epsilon(\zeta^-(\epsilon))} = 1 - \frac{i\pi}{f(z)}\epsilon - \frac{\pi^2}{2f(z)^2}\epsilon^2 + O(\epsilon^3).
$$

As in the previous lemma the constant in the O depends on K.

Proof. The invariance of the parabolic curves gives

$$
\begin{aligned}
\frac{f_\epsilon(z) - f_\epsilon(\zeta^+(\epsilon))}{f_\epsilon(z) - f_\epsilon(\zeta^-(\epsilon))} = \frac{f_\epsilon(z) - \zeta^+(\tilde{g}(\epsilon))}{f_\epsilon(z) - \zeta^-(\tilde{g}(\epsilon))} &= 1 - \frac{i\pi}{f_\epsilon(z)}\tilde{g}(\epsilon) + \frac{\pi^2}{2f_\epsilon(z)^2}\tilde{g}(\epsilon^2) + O(\epsilon^3) \\
&= 1 - \frac{i\pi}{f(z)}\epsilon + \frac{\pi^2}{2f(z)^2}\epsilon^2 + O(\epsilon^3).
\end{aligned}
$$

The last equality uses the fact that $\tilde{g}(\epsilon) = \epsilon + O(\epsilon^3)$. □

Proposition 5.8. *There exists a constant $C_0 \in \mathbb{C}$ (depending only on f and g) such that*

$$A(\epsilon, z) = \epsilon A_0(z) + \epsilon^3 C_0 + O(\epsilon^4, \epsilon^3 z),$$

where the constants in the O are uniform for $(z, \epsilon) \in \mathbb{C}^2$ near $(0,0)$ (with $\mathrm{Re}(\epsilon) > 0$).

Proof. Let K be a compact of $\mathbb{C}^*$. Then by the two previous lemmas, we have

$$\begin{aligned}
A(\epsilon, z) &= \frac{1}{i\pi} \log\left(\frac{z - \zeta^-(\epsilon)}{z - \zeta^+(\epsilon)} \cdot \frac{f_\epsilon(z) - f_\epsilon(\zeta^+(\epsilon))}{f_\epsilon(z) - f_\epsilon(\zeta^-(\epsilon))}\right) - \epsilon \\
&= -\frac{1}{i\pi} \log\left(1 - \frac{i\pi}{z}\epsilon - \frac{\pi^2}{2z^2}\epsilon^2 + O(\epsilon^3)\right) + \frac{1}{i\pi} \log\left(1 - \frac{i\pi}{f(z)}\epsilon - \frac{\pi^2}{2f(z)^2}\epsilon^2 + O(\epsilon^3)\right) - \epsilon \\
&= \frac{\epsilon}{z} - \frac{\epsilon}{f(z)} - \epsilon + O(\epsilon^3) = \epsilon A_0(z) + O(\epsilon^3).
\end{aligned}$$

Here the constant in the O still depends on $K \subset \mathbb{C}^*$. Let $\phi_\epsilon(z) := (A(\epsilon, z) - \epsilon A_0(z))/\epsilon^3$. By Proposition 5.5, ϕ_ϵ is holomorphic on $\mathbb{D}(0, r)$. We have proved that for all compact $K \subset \mathbb{C}^*$, for all $z \in K$, and for all small $\epsilon \neq 0$ with $\mathrm{Re}(\epsilon) > 0$, we have $|\phi_\epsilon(z)| \leq C_K$. By taking $K = \left\{|z| = \frac{1}{2}r\right\}$ we therefore obtain the same estimate $|\phi_\epsilon(z)| \leq C_K$ for all $|z| \leq \frac{1}{2}r$ because of the maximum modulus principle. This gives the desired uniformity. $\square$

Lemma 5.9. *If $\zeta^\pm(\epsilon) = \pm\frac{\pi}{2}i\epsilon + c_2\epsilon^2 + c_3^\pm\epsilon^3 + O(\epsilon^4)$ and $f(z) = z + z^2 + z^3 + bz^4 + O(z^5)$, then*

$$C_0 = \frac{-3b\pi^3 + 2\pi^3 + 12c_2\pi + 12i(c_3^- - c_3^+)}{12\pi}.$$

Proof. By repeating the computations from Lemmas 5.6 and 5.7 with one additional order of significance, one obtains

$$\begin{aligned}
\frac{z - \zeta^+(\epsilon)}{z - \zeta^-(\epsilon)} &= 1 - \frac{i\pi}{z}\epsilon - \frac{\pi^2}{2z^2}\epsilon^2 + \left(\frac{c_3^- - c_3^+}{z} - \frac{i\pi c_2}{z^2} + \frac{i\pi^3}{4z^3}\right)\epsilon^3 + O(\epsilon^3), \\
\frac{f_\epsilon(z) - f_\epsilon(\zeta^+(\epsilon))}{f_\epsilon(z) - f_\epsilon(\zeta^-(\epsilon))} &= 1 - \frac{i\pi}{f(z)}\epsilon - \frac{\pi^2}{2f^2(z)}\epsilon^2 + \left(\frac{i\pi^3}{4f^2(z)} + \frac{c_3^- - c_3^+}{f(z)} - \frac{i\pi c_2}{f^2(z)} + \frac{i\pi^3}{4f^3(z)}\right)\epsilon^3 + O(\epsilon^4).
\end{aligned}$$

Plugging these two equations into the formula for $A(\epsilon, z)$, and using the power series expansions of $1/f(z)^j$ for $j = 1, \ldots, 3$, one notices again that all terms involving negative powers of z cancel, either by the argument used in the proof of the previous proposition, or by lengthy computations using

$$c_2 = \frac{\pi^2}{8} - \frac{1}{4}.$$

Summing the terms that do not depend on z gives the desired result. $\square$

Lemma 5.10. *We have*

$$c_3^+ = -c_3^- = \frac{1}{i\pi}\left(\frac{3}{16} + \frac{5\pi^4}{64} - \frac{b\pi^4}{16} - \frac{\pi^2}{4}\right).$$

We will omit the proof, which is a long but direct computation, starting from the functional equation $f_\epsilon \circ \zeta^\pm(\epsilon) = \zeta^\pm \circ \tilde{g}(\epsilon)$ and identifying coefficients in powers of ϵ.

In particular, it follows that

$$\begin{aligned} C_0 &= -\frac{b\pi^2}{4} - \frac{1}{4} + \frac{7\pi^2}{24} + \Big(\frac{b\pi^2}{8} + \frac{1}{2} - \frac{5\pi^2}{32} - \frac{3}{8\pi^2}\Big) \\ &= \frac{-b\pi^2}{8} + \frac{13\pi^2}{96} - \frac{3}{8\pi^2} + \frac{1}{4}. \end{aligned}$$

5B. *Convergence result.* For the rest of Section 5 we fix a compact subset $K \times K' \subset \mathcal{B}_f \times \mathcal{B}_g$ and a point $(z_0, w_0) \in K \times K'$. Moreover we assume that n is sufficiently large so that $g^{n^2}(K')$ is contained in a petal $\mathcal{P}$ from Proposition 4.1. Unless otherwise stated, all the constants appearing in estimates depend only on the compact $K \times K'$, but not on the point (z_0, w_0) nor the integer n.

Let $f_j(z) := f(z) + \frac{\pi^2}{4} w_{n^2+j}$, where $w_{n^2+j} := g^{n^2+j}(w_0)$. Let $z_j := f_j \circ f_{j-1} \circ \cdots \circ f_1(z_0)$. Let $F_{m,p} := f_m \circ \cdots \circ f_{p+1}$, and let $\epsilon_j := \sqrt{w_{n^2+j}}$.

The strategy of the proof of Theorem 5.33 is as follows: we will use *approximate Fatou coordinates* $\phi_n^{\iota/o}$ and prove that on some appropriate domains $\phi_n^{\iota/o}$ converges locally uniformly to $\phi_f^{\iota/o}$ (with a known error term of order $1/n$). Moreover, we will compute $\phi_n^\iota(z_0)$ and $\phi_n^o(z_{2n+1})$, again at a precision of order $1/n$. This will allow us to compare accurately z_{2n+1} and $\mathcal{L}_f(z_0) = (\phi_f^o)^{-1} \circ \phi_f^\iota(z_0)$. This approach differs from [Astorg et al. 2016] in that approximate Fatou coordinates in that work were only used at small scale near 0, while here they are defined on a whole petal: this simplifies the comparison with the actual Fatou coordinates $\phi_f^{\iota/o}$. The approach used here is strongly inspired by [Bedford et al. 2017].

Definition 5.11. Let

$$Z_j^{i/o} := \psi_{\epsilon_j}^{\iota/o}(z_j) = \frac{1}{i\pi} \log \frac{\zeta^+(\epsilon_j) - z_j}{z_j - \zeta^-(\epsilon_j)} \pm 1.$$

Observe that by definition of $A(\epsilon, z)$,

$$A(\epsilon_j, z_j) = Z_{j+1}^i - Z_j^i - \epsilon_j. \tag{14}$$

Proposition 5.12. *We have*

$$\psi_{\epsilon_j}^{\iota/o}(z_0) = -\frac{\epsilon_j}{z_0} + O\Big(\frac{1}{n^3}\Big).$$

Proof. This follows from computations similar to those appearing in the proof of Proposition 5.5 (recall as well that $\epsilon_j = O(1/n)$). ☐

We now introduce approximate incoming Fatou coordinates:

Definition 5.13. Let

$$\phi_n^\iota(z_0) := \frac{1}{\epsilon_n} Z_n - \frac{1}{\epsilon_n} \sum_{j=1}^{n-1} \epsilon_j.$$

Let D_ϵ be the disk of radius $\frac{1}{2}|\zeta^+(\epsilon) - \zeta^-(\epsilon)|$ centered at $\frac{1}{2}(\zeta^+(\epsilon) + \zeta^-(\epsilon))$. Let $\mathcal{S}(\epsilon, r)$ be the union of the two disks of radius r that both contain the points $\zeta^+(\epsilon), \zeta^-(\epsilon)$ on their boundary. Here r will be a sufficiently small number, to be fixed in the paragraph before Lemma 5.14. The definition of $\mathcal{S}(\epsilon, r)$ of course only makes sense when the distance between $\zeta^+(\epsilon)$ and $\zeta^-(\epsilon)$ is less than $2r$, which once r is fixed will be satisfied for ϵ sufficiently small. We note that the choice of r will depend on the map f, but not on ϵ.

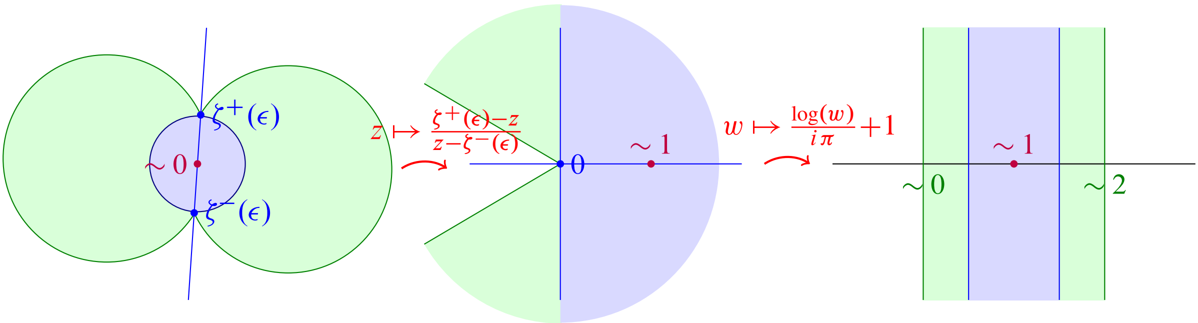

Figure 1. *The sets $D_\epsilon \subset S(\epsilon, r)$ and their images under ψ_ϵ^ι.*

The line L_ϵ through $\zeta^+(\epsilon)$ and $\zeta^-(\epsilon)$ cuts the complex plane into the left half-plane H_ϵ^ι and the right half-plane H_ϵ^o. We define $\mathcal{S}^{\iota/o}(\epsilon, r) := \mathcal{S}(\epsilon, r) \cap H_\epsilon^{\iota/o}$. The map ψ_ϵ^ι maps the disk D_ϵ to the strip $\left[\frac{1}{2}, \frac{3}{2}\right] \times i\mathbb{R}$. The image of $\mathcal{S}^\iota(\epsilon, r)$ is bounded by two vertical lines, intersecting the real line in a point of the form $0 + O(\epsilon)$ and in the point 1; see Figure 1. In particular we can find $0 < \alpha < \beta$ such that

$$[\beta\epsilon, 1] \times i\mathbb{R} \subset \psi_\epsilon^\iota(\mathcal{S}^\iota(\epsilon, r)) \subset [\alpha\epsilon, 1] \times i\mathbb{R}$$

for all ϵ. We define $\mathcal{S}^\iota(0, r) := \mathbb{D}(-r, r)$.

Recall that $A(\epsilon, \cdot)$ is analytic on a small disk $\mathbb{D}(0, R)$ centered at the origin. Moreover there exists $R > 0$ such that $|A(\epsilon, z)| \leq \frac{1}{12}|\epsilon|$ for all $z \in \mathbb{D}(0, R)$ and ϵ in the petal $\mathcal{P}_\delta$ defined in Section 4. By taking smaller R if necessary we my assume that f is 1-Lipschitz on $\mathbb{D}(-R, R)$.

Now let us assume that $r \ll R$ is sufficiently small so that $\mathcal{S}(\epsilon, r) \subset \mathbb{D}(0, R)$ for all $\epsilon > 0$, and note that for every compact set $K \subset \mathcal{B}_f$ there exist $n', \epsilon' > 0$ so that $f^n(K) \subset \mathcal{S}^\iota(\epsilon, r) \cap \mathbb{D}(-R, R)$ for all $n \geq n'$ and all $\epsilon \geq \epsilon'$. We now fix this r.

Lemma 5.14. *Let $K \times K' \subset \mathcal{B}_f \times \mathcal{B}_g$ be a compact set. There exist $n_0, m_0 > 0$ such that for all $(z_0, w_0) \in K \times K'$ and all $n > n_0$ we have*

(1) $z_j \in \mathcal{S}^\iota(\epsilon_j, r) \cup D_{\epsilon_j}$ *for all* $m_0 \leq j \leq n-1$,

(2) $z_j \in D_{\epsilon_j}$ *for all* $\frac{2}{3}n \leq j \leq n-1$,

(3) *If* $z_k \in \mathcal{S}(\epsilon_k, r)$ *for all* $m_0 \leq k \leq j$, *then* $|\mathrm{Im}(\psi_{\epsilon_{j+1}}^\iota(z_{j+1}))| < 1$,

where $\epsilon_j := \sqrt{w_{n^2+j}} = 1/n + O(j/n^3)$.

Proof. There exists $m_0 > 0$ so that $f^{m_0}(K) \subset S^\iota(0, r)$. Let n_0 be sufficiently large so that for all $(z_0, w_0) \in K \times K'$ we have:

(i) $\beta|\epsilon_{m_0}| < \mathrm{Re}(\psi_{\epsilon_{m_0}}^\iota(z_{m_0})) < \frac{1}{6}$.

(ii) $|\mathrm{Im}(\psi_{\epsilon_{m_0}}^\iota(z_{m_0}))| < \frac{1}{2}$.

(iii) $|\epsilon_j| < 2\,\mathrm{Re}(\epsilon_j) < R$ for $0 \leq j \leq 2n+1$.

Indeed, (i) and (ii) follow from the equality $\psi_\epsilon^\iota(z) = -\epsilon/z + O(\epsilon^3)$ and (iii) follows from the fact that $\epsilon_j = 1/n + O(j/n^3)$. Note that the constants in O depend only on the compact $K \times K'$ and not on n or j.

Recall that by our assumption $\mathcal{S}^\iota(\epsilon_j, r) \subset \mathbb{D}(0, R)$ for all j and that $|A(\epsilon_j, z)| \leq \frac{1}{12}|\epsilon_j|$ for all $z \in \mathbb{D}(0, R)$. By (i) we have $z_{m_0} \in \mathcal{S}^\iota(\epsilon_{m_0}, r)$ and observe that for $z_j \in \mathcal{S}^\iota(\epsilon_j, r)$ we have

$$\frac{5}{6}\operatorname{Re}(\epsilon_j) < \operatorname{Re}(\psi^\iota_{\epsilon_{j+1}}(z_{j+1}) - \psi^\iota_{\epsilon_j}(z_j)) < \frac{7}{6}\operatorname{Re}(\epsilon_j).$$

It follows that

$$\beta|\epsilon_{m_0}| + \frac{5}{6}\sum_{k=m_0}^{j-1}\operatorname{Re}(\epsilon_k) < \operatorname{Re}(\psi^\iota_{\epsilon_j}(z_j)) < \frac{1}{6} + \frac{7}{6}\sum_{k=m_0}^{j-1}\operatorname{Re}(\epsilon_k), \tag{15}$$

and since $\operatorname{Re}(\epsilon_k) = 1/n + O(k/n^3)$ we have

$$\beta|\epsilon_{m_0}| < \operatorname{Re}(\psi^\iota_{\epsilon_j}(z_j)) < \frac{3}{2}$$

for all $m_0 \leq j \leq n-1$ as long as n is sufficiently large. This proves (1).

For (2) observe that

$$-\frac{1}{2} < -1 + \frac{5}{6}\sum_{k=0}^{j-1}\operatorname{Re}(\epsilon_k)$$

for all $\frac{2}{3}n \leq j \leq n-1$ as long as n is sufficiently large.

Finally for (3) observe that (15) implies that $z_k \in \mathcal{S}(\epsilon_k, r)$ for all $0 \leq k \leq j$ can only hold for some $j < 3n$. By (ii) and (iii) we have

$$\operatorname{Im}(\epsilon_j) - \frac{1}{6}\operatorname{Re}(\epsilon_j) < \operatorname{Im}(\psi^\iota_{\epsilon_{j+1}}(z_{j+1}) - \psi^\iota_{\epsilon_j}(z_j)) < \operatorname{Im}(\epsilon_j) + \frac{1}{6}\operatorname{Re}(\epsilon_j)$$

for $z_j \in \mathcal{S}^\iota(\epsilon_j, r)$; hence

$$-\frac{1}{2} + \sum_{k=m_0}^{j-1}\operatorname{Im}(\epsilon_k) - \frac{1}{6}\operatorname{Re}(\epsilon_k) < \operatorname{Im}(\psi_{\epsilon_j}(z_j)) < \frac{1}{2} + \sum_{k=m_0}^{j-1}\operatorname{Im}(\epsilon_k) + \frac{1}{6}\operatorname{Re}(\epsilon_k).$$

Since $\operatorname{Im}(\epsilon_k) = O(k/n^3)$ we can conclude that (3) holds as long as n is sufficiently large. $\square$

Lemma 5.15. *For $0 \leq j \leq n-1$ we have*

$$z_j - f^j(z_0) = O\left(\frac{j}{n^2}\right).$$

Proof. Let m_0 be as in Lemma 5.14. Since m_0 is independent from n it is easy to see that for all $0 \leq j \leq m_0$ we have

$$z_j - f^j(z_0) = \frac{\pi^2}{4}\sum_{k=0}^{j-1}\left(\epsilon_k^2 + O\left(\frac{1}{n^3}\right)\right) = O\left(\frac{j}{n^2}\right).$$

Let $V_\epsilon := \{z \in \mathcal{S}^\iota(\epsilon, r) : |\operatorname{Im}(\psi^\iota_\epsilon(z))| < 1\}$ and observe that $V_\epsilon \backslash D_\epsilon \subset \mathbb{D}(-R, R)$ for all sufficiently small ϵ. Since by our assumption f is 1-Lipschitz on $\mathbb{D}(-R, R)$, it follows by (1) in Lemma 5.14 that for all $m_0 < j < \frac{2}{3}n$ we have

$$|z_j - f^j(z_0)| < |z_{j-1} - f^{j-1}(z_0)| + \frac{\pi^2}{4}\epsilon_{j-1}^2$$
$$< |z_{m_0} - f^{m_0}(z_0)| + \frac{\pi^2}{4}\sum_{k=m_0}^{j-1}\epsilon_k^2 = \frac{\pi^2}{4}\sum_{k=0}^{j-1}\left(\frac{1}{n^2} + O\left(\frac{1}{n^3}\right)\right) = O\left(\frac{j}{n^2}\right).$$

Finally for $\frac{2}{3}n \le j \le n-1$, by item (2) of Lemma 5.14, we have $z_j \in D_{\epsilon_j}$, and in particular $z_j = O(1/n)$. It follows that

$$z_j - f^j(z_0) = O\left(\frac{1}{n}, \frac{1}{j}\right) = O\left(\frac{j}{n^2}\right),$$

where the last equality follows from the fact that $\frac{2}{3}n \le j \le n-1$. □

Lemma 5.16. *We have*

$$\sum_{j=0}^{n-1} A_0(z_j) - A_0(f^j(z_0)) = (b-1)\sum_{j=0}^{n-1} z_j^2 - f^j(z_0)^2 + O\left(\frac{\log n}{n^2}\right).$$

Proof. Recall that $f(z) = z + z^2 + z^3 + bz^4 + O(z^5)$ and $A_0(z) = -1/f(z) + 1/z - 1$. An elementary computation gives $A_0(0) = A_0'(0) = 0$ and $A_0''(0) = 2(b-1)$.

To simplify the notation, let $y_j := f^j(z_0)$. We have

$$A_0(z_j) - A_0(y_j) = A_0'(y_j)(z_j - y_j) + \tfrac{1}{2}A_0''(y_j)(z_j - y_j)^2 + O((z_j - y_j)^3)$$
$$= (y_j A_0''(0) + O(y_j^2))(z_j - y_j) + \tfrac{1}{2}(A_0''(0) + O(y_j))(z_j - y_j)^2 + O((z_j - y_j)^3).$$

By Lemma 5.15 we have $z_j - y_j = O(j/n^2)$; hence

$$A_0(z_j) - A_0(y_j) = y_j A_0''(0)(z_j - y_j) + \tfrac{1}{2}A_0''(0)(z_j - y_j)^2 + O\left(\frac{1}{jn^2}, \frac{j}{n^4}, \frac{j^3}{n^6}\right)$$
$$= (b-1)(2y_j(z_j - y_j) + (z_j - y_j)^2) + O\left(\frac{1}{jn^2}, \frac{j}{n^4}, \frac{j^3}{n^6}\right)$$
$$= (b-1)(z_j^2 - y_j^2) + O\left(\frac{1}{jn^2}, \frac{j}{n^4}, \frac{j^3}{n^6}\right).$$

It follows that

$$\sum_{j=0}^{n-1} A_0(z_j) - A_0(y_j) = (b-1)\sum_{j=0}^{n-1} z_j^2 - y_j^2 + O\left(\frac{\log n}{n^2}\right).$$ □

Lemma 5.17. *For* $0 \le j \le n-1$, *let*

$$\gamma_j := j + \sum_{k=0}^{j-1} A_0(f^k(z_0)) \quad \textit{and} \quad x_j := \sum_{k=0}^{j-1} \epsilon_k + A(\epsilon_k, z_k). \tag{16}$$

Then

$$x_j = \frac{\gamma_j}{n} + O\left(\frac{j^2}{n^3}\right) = \frac{j}{n} + O\left(\frac{1}{n}\right).$$

In particular, there exists $k \in \mathbb{N}$ independent from n such that for all $k \leq j \leq n-k$

$$\alpha_j := \cot\left(\frac{\pi}{2}x_j\right)$$

is well-defined and strictly positive.

Proof. According to Lemma 5.15, for $0 \leq j \leq n-1$ we have that $z_j - f^j(z_0) = O(j/n^2)$. In particular, $z_j = O(1)$. By Proposition 5.8, we have for every $0 \leq k \leq n-1$

$$A(\epsilon_k, z_k) = \epsilon_k A_0(z_k) + O(\epsilon_k^3, z_k\epsilon_k^3) = \epsilon_k A_0(z_k) + O\left(\frac{1}{n^3}\right) \tag{17}$$

(indeed, by Lemma 5.15, $z_k = f^k(z_0) + O(k/n^2)$, so in particular $z_k = O(1)$). By Lemma 5.2, $\epsilon_k = 1/n + O(k/n^3)$; hence

$$\begin{aligned} x_j &= \sum_{k=0}^{j-1} \epsilon_k + A(\epsilon_k, z_k) = \sum_{k=0}^{j-1} \frac{1}{n} + \frac{1}{n}A_0(z_k) + O\left(\frac{k}{n^3}\right) \\ &= \frac{j}{n} + \frac{1}{n}\sum_{k=0}^{j-1} A_0(f^k(z_0)) + O\left(\frac{k}{n^2}, A_0'(f^k(z_0))(z_k - f^k(z_0))\right) \\ &= \frac{j}{n} + \frac{1}{n}\sum_{k=0}^{j-1} A_0(f^k(z_0)) + O\left(\frac{k}{n^2}, \frac{1}{k}\cdot\frac{k}{n^2}\right) = \frac{\gamma_j}{n} + O\left(\frac{j^2}{n^3}\right). \end{aligned}$$

Since $\gamma_j = j + O(1)$, we also have

$$\frac{\gamma_j}{n} = \frac{j}{n} + O\left(\frac{1}{n}\right).$$

Finally, the last assertion follows from the preceding equality and the fact that, for $x \in \left(0, \frac{\pi}{2}\right)$, $\cot(x) > 0$. ☐

Lemma 5.18. *Let*

$$u(x) := \frac{2}{\pi}\tan\left(\frac{\pi}{2}x\right), \quad \Phi(x) := \frac{x^2 - u(x)^2}{x^2u(x)^2} \quad \textit{and} \quad \beta_j := \frac{2n}{\pi\alpha_j}.$$

We have

$$\frac{\gamma_j^2 - \beta_j^2}{\gamma_j^2\beta_j^2} = \frac{1}{n^2}\Phi(x_j) + O\left(\frac{1}{jn^2}\right).$$

Proof. We have

$$\frac{\gamma_j^2 - \beta_j^2}{\gamma_j^2\beta_j^2} = \frac{n^2}{n^4}\frac{(\gamma_j^2/n^2) - u(x_j)^2}{u(x_j)^2(\gamma_j^2/n^2)}.$$

Now recall that by Lemma 5.17, $\gamma_j/n = x_j + O(j^2/n^3)$, so that $\gamma_j^2/n^2 = x_j^2 + O(j^3/n^4)$. So

$$\frac{\gamma_j^2 - \beta_j^2}{\gamma_j^2\beta_j^2} = \frac{1}{n^2}\frac{x_j^2 - u(x_j)^2 + O(j^3/n^4)}{u(x_j^2)(x_j^2 + O(j^3/n^4))} = \frac{1}{n^2}\frac{x_j^2 - u(x_j)^2}{u(x_j^2)(x_j^2 + O(j^3/n^4))} + O\left(\frac{j^3}{x_j^2u(x_j)^2n^6}\right)$$

and note that

$$\frac{1}{x_j^2 u(x_j)^2} = O\left(\frac{1}{x_j^4}\right) = O\left(\frac{n^4}{j^4}\right).$$

Therefore

$$\begin{aligned}\frac{\gamma_j^2-\beta_j^2}{\gamma_j^2\beta_j^2} &= \frac{1}{n^2}\frac{x_j^2-u(x_j)^2}{u(x_j^2)(x_j^2+O(j^3/n^4))} + O\left(\frac{1}{jn^2}\right)\\ &= \frac{1}{n^2}\frac{x_j^2-u(x_j)^2}{u(x_j^2)x_j^2(1+O(j^3/(n^4x_j^2)))} + O\left(\frac{1}{jn^2}\right)\\ &= \frac{\Phi(x_j)}{n^2}\left(1+O\left(\frac{j}{n^2}\right)\right) + O\left(\frac{1}{jn^2}\right)\\ &= \frac{\Phi(x_j)}{n^2} + O\left(\frac{1}{jn^2}\right).\end{aligned}$$

Note that in the last line, we used the fact that Φ has only removable singularities at $x=0$ and $x=1$, so that $\Phi(x_j)=O(1)$. □

Proposition 5.19. *There exists a universal constant $C_1 \in \mathbb{R}$ such that*

$$\sum_{j=0}^{n-1} A_0(z_j) - A_0(f^j(z)) = \frac{C_1(b-1)}{n} + O\left(\frac{\log n}{n^2}\right).$$

More precisely, $C_1 := \int_0^1 \Phi(x)\,dx = \frac{1}{4}(4-\pi^2)$.

Proof. We have, for $0 \le j \le n-1$,

$$z_j = \psi_{\epsilon_j}^{-1}(Z_j^\iota) = -\frac{\pi}{2n}\cot\left(\frac{\pi}{2}Z_j^\iota\right) + O\left(\frac{1}{n^2}\right)$$

and

$$Z_j^\iota = Z_0^\iota + \sum_{k=0}^{j-1} \epsilon_k + A(\epsilon_k, z_k).$$

Recalling the notation $x_j := \sum_{k=0}^{j-1} \epsilon_k + A(\epsilon_k, z_k)$ and $\alpha_j := \cot((\pi/2)x_j)$ from Lemma 5.17, and using the trigonometric formula

$$\cot(a+b) = \frac{\cot a \cot b - 1}{\cot a + \cot b},$$

we therefore obtain

$$z_j = -\frac{\pi}{2n}\frac{\cot((\pi/2)Z_0^\iota)\alpha_j - 1}{\alpha_j + \cot((\pi/2)Z_0^\iota)} + O\left(\frac{1}{n^2}\right). \tag{18}$$

Let k be as in Lemma 5.17, so that $\alpha_j > 0$ for $k \le j \le n-k$. We have

$$\cot\left(\frac{\pi}{2}Z_0^\iota\right) = -\frac{2n}{\pi}z_0 + O\left(\frac{1}{n}\right),$$

so that

$$z_j = -\frac{\pi}{2n}\frac{\cot((\pi/2)Z_0^\iota)\alpha_j}{\alpha_j + \cot((\pi/2)Z_0^\iota)} + O\left(\frac{1}{n^2}\right) = \frac{z_0\alpha_j}{\alpha_j - z_0(2n/\pi)} + O\left(\frac{1}{n^2}\right). \tag{19}$$

Finally, with $\beta_j := 2n/(\pi\alpha_j)$, we get

$$z_j = -\frac{1}{-1/z_0 + \beta_j} + O\left(\frac{1}{n^2}\right). \tag{20}$$

On the other hand, from the definition of A_0 it follows that $\sum_{k=0}^{j-1} A_0(f^k(z_0)) = 1/z_0 - 1/f^j(z_0) - j$, which we may rewrite as

$$f^j(z_0) = -\frac{1}{-1/z_0 + \gamma_j}. \tag{21}$$

Therefore

$$z_j - f^j(z_0) = \frac{\beta_j - \gamma_j}{(-1/z_0 + \beta_j)(-1/z_0 + \gamma_j)} + O\left(\frac{1}{n^2}\right). \tag{22}$$

Now note that $j = O(\beta_j)$; indeed,

$$\frac{1}{\beta_j} = O\left(\frac{\cot(j\pi/n + O(1/n))}{n}\right) = O\left(\frac{n/j}{n}\right) = O\left(\frac{1}{j}\right).$$

Therefore

$$\begin{aligned}\frac{1}{(-1/z_0 + \gamma_j)(-1/z_0 + \beta_j)} &= \frac{1}{(\gamma_j + O(1))(\beta_j + O(1))} \\ &= \frac{1}{\gamma_j\beta_j + O(\beta_j)} = \frac{1}{\gamma_j\beta_j} + O\left(\frac{1}{\gamma_j^2\beta_j}\right).\end{aligned}$$

Thus, setting $y_j := f^j(z_0)$,

$$\begin{aligned}z_j^2 - y_j^2 &= (z_j - y_j)(z_j + y_j) \\ &= \left(\frac{\beta_j - \gamma_j}{\beta_j\gamma_j} + O\left(\frac{\beta_j - \gamma_j}{\gamma_j^2\beta_j}\right)\right)\left(-\frac{\beta_j + \gamma_j}{\beta_j\gamma_j} + O\left(\frac{\beta_j + \gamma_j}{\gamma_j^2\beta_j}\right)\right) \\ &= \frac{\gamma_j^2 - \beta_j^2}{\beta_j^2\gamma_j^2} + O\left(\frac{\beta_j^2 - \gamma_j^2}{\beta_j^2\gamma_j^3}\right) \\ &= \frac{1}{n^2}\Phi(x_j) + O\left(\frac{1}{jn^2}\right) \quad \text{by Lemma 5.18.}\end{aligned}$$

Therefore by Lemma 5.16,

$$\begin{aligned}\sum_{j=0}^{n-1} A_0(z_j) - A_0(y_j) &= \left(\frac{b-1}{n^2}\sum_{j=0}^{n-1}\Phi(x_j)\right) + O\left(\frac{\log n}{n^2}\right) \\ &= \frac{b-1}{n}\int_0^1 \Phi(x)\,dx + O\left(\frac{\log n}{n^2}\right).\end{aligned}$$

In the last equality, we recognize a Riemann sum with subdivision $(x_j)_{0\leq j\leq n-1}$. Finally, we have

$$\int_0^1 \Phi(x)\,dx = \frac{\pi}{2}\int_0^{\frac{\pi}{2}} \cot^2 t - \frac{1}{t^2}\,dt = -\frac{\pi}{2}\Big[\Big(\cot t - \frac{1}{t}\Big) + t\Big]_0^{\frac{\pi}{2}} = 1 - \frac{\pi^2}{4}. \qquad \square$$

5B1. *Incoming part.* The following error estimate is one of the two crucial estimates that we will obtain in this section: it measures accurately how close ϕ_n^ι is to the incoming Fatou coordinate ϕ_f^ι. This estimate differs from those obtained in [Astorg et al. 2016] in that we compare ϕ_n^ι with ϕ_f^ι on a definite region of $\mathcal{B}_f$ (independent from n), instead of comparing the two at small scale near the origin, compare with [Astorg et al. 2016, Property 1, p. 10]. Moreover, the point of Proposition 5.20 is to push the precision of the estimate further and obtain the first error term $E^\iota(z_0)/n$, which cannot be easily obtained from the computations in [Astorg et al. 2016].

Proposition 5.20. *We have*

$$\phi_n^\iota(z_0) = \phi_f^\iota(z_0) + \frac{E^\iota(z_0)}{n} + O\Big(\frac{\log n}{n^2}\Big),$$

where $E^\iota(z) := C_0 + (C_1 - 1)(b-1) + \frac{1}{2}\phi^\iota(z_0)$.

Proof. Recall that by definition,

$$\phi_f^\iota(z) = \lim_{n\to\infty} -\frac{1}{f^n(z)} - n = \lim_{n\to\infty} -\frac{1}{z} + \sum_{j=0}^{n-1} A_0(f^j(z)).$$

Similarly, we have

$$\sum_{j=0}^{n-1} A(\epsilon_j, z_j) = \sum_{j=0}^{n-1} Z_{j+1}^\iota - Z_j^\iota - \epsilon_j = Z_n^\iota - Z_0^\iota - \sum_{j=0}^{n-1}\epsilon_j,$$

and thus

$$\phi_n^\iota(z_0) = \frac{Z_n^\iota}{\epsilon_n} - \frac{1}{\epsilon_n}\sum_{j=0}^{n-1}\epsilon_j = \frac{Z_0^\iota}{\epsilon_n} + \frac{1}{\epsilon_n}\sum_{j=0}^{n-1} A(\epsilon_j, z_j).$$

Therefore

$$\phi_n^\iota(z_0) - \phi_f^\iota(z_0) = E_1 + E_2 + E_3, \tag{23}$$

where

$$E_1 := \frac{Z_0^\iota}{\epsilon_n} + \frac{1}{z_0},$$

$$E_2 := \frac{1}{\epsilon_n}\sum_{j=0}^{n-1} A(\epsilon_j, z_j) - \sum_{j=0}^{n-1} A_0(f^j(z_0)),$$

$$E_3 := -\sum_{j=n}^{\infty} A_0(f^j(z_0)).$$

We will now estimate each of the error terms E_i separately. For $j \in \mathbb{N}$, we set $y_j := f^j(z_0)$.

Lemma 5.21. *We have*

$$E_1 = -\frac{1}{2nz_0} + O\Big(\frac{1}{n^2}\Big).$$

Proof of lemma. We have

$$\begin{aligned}\frac{Z_0^\iota}{\epsilon_n} &= \frac{1}{\epsilon_n}\psi_{\epsilon_0}(z_0) = -\frac{\epsilon_0}{\epsilon_n z_0} + O\left(\frac{\epsilon_0^3}{\epsilon_n}\right) \quad \text{(by Proposition 5.12)}\\ &= -\frac{1}{z_0}\sqrt{\frac{n^2+n+O(1)}{n^2+O(1)}} + O\left(\frac{1}{n^2}\right)\\ &= -\frac{1}{z_0} - \frac{1}{2nz_0} + O\left(\frac{1}{n^2}\right).\end{aligned}$$

□

Lemma 5.22. *We have*

$$E_2 = \frac{1}{n}\left(\frac{1}{2z_0} + \frac{1}{2}\phi_f^\iota(z_0) + C_0 + C_1(b-1)\right) + O\left(\frac{\log n}{n^2}\right).$$

Proof of lemma. Recall that we have

$$A(\epsilon, z) = \epsilon A_0(z) + C_0\epsilon^3 + O(z\epsilon^3, \epsilon^4),$$

so that

$$\begin{aligned}E_2 &= \frac{1}{\epsilon_n}\sum_{j=0}^{n-1} A(\epsilon_j, z_j) - \sum_{j=0}^{n-1} A_0(y_j)\\ &= \frac{1}{\epsilon_n}\sum_{j=0}^{n-1} \epsilon_j A_0(z_j) + C_0\epsilon_j^3 + O\left(\frac{z_j}{n^3}\right) - \epsilon_n A_0(y_j).\end{aligned}$$

Therefore

$$E_2 = \left(\sum_{j=0}^{n-1} \frac{\epsilon_j}{\epsilon_n} A_0(z_j) - A_0(y_j)\right) + \left(\sum_{j=0}^{n-1} C_0\frac{\epsilon_j^3}{\epsilon_n} + O\left(\frac{z_j}{n^2}\right)\right)$$

and

$$\sum_{j=0}^{n-1} C_0\frac{\epsilon_j^3}{\epsilon_n} + O\left(\frac{z_j}{n^2}\right) = C_0\sum_{j=0}^{n-1}\frac{1}{n^2} + O\left(\frac{1}{n^2 j}\right) = \frac{C_0}{n} + O\left(\frac{\log n}{n^2}\right).$$

On the other hand, we have

$$\sum_{j=0}^{n-1} \frac{\epsilon_j}{\epsilon_n} A_0(z_j) - A_0(y_j) = \sum_{j=0}^{n-1}\left(\frac{\epsilon_j}{\epsilon_n} - 1\right) A_0(z_j) + \sum_{j=0}^{n-1} A_0(z_j) - A_0(y_j). \tag{24}$$

Now note that

$$\sum_{j=0}^{n-1}\left(\frac{\epsilon_j}{\epsilon_n} - 1\right) A_0(z_j) = \sum_{j=0}^{n-1}\left(\frac{\epsilon_j}{\epsilon_n} - 1\right) A_0(y_j) + \sum_{j=0}^{n-1}\left(\frac{\epsilon_j}{\epsilon_n} - 1\right)(A_0(z_j) - A_0(y_j)),$$

and that

$$\left|\sum_{j=0}^{n-1}\left(\frac{\epsilon_j}{\epsilon_n} - 1\right)(A_0(z_j) - A_0(y_j))\right| \le \max_{0\le j\le n-1}\left|1 - \frac{\epsilon_j}{\epsilon_n}\right| \cdot \sum_{j=0}^{n-1}|A_0(z_j) - A_0(y_j)| = O\left(\frac{1}{n^2}\right),$$

by Lemma 5.2 and Proposition 5.19. Another consequence of Lemma 5.2 is that

$$\frac{\epsilon_j}{\epsilon_n} - 1 = \frac{1}{2n}\left(1 - \frac{j}{n} + O\left(\frac{1}{n}\right)\right). \tag{25}$$

Therefore, by (24), (25) and Proposition 5.19,

$$\begin{aligned}\sum_{j=0}^{n-1} \frac{\epsilon_j}{\epsilon_n} A_0(z_j) - A_0(y_j) &= \frac{C_1(b-1)}{n} + \frac{1}{2n}\sum_{j=0}^{n-1} A_0(y_j) + O\left(\frac{\log n}{n^2}\right)\\ &= \frac{C_1(b-1)}{n} + \left(\frac{1/z_0 + \phi_f^\iota(z_0)}{2n}\right) + O\left(\frac{\log n}{n^2}\right).\end{aligned}$$

Therefore, as announced, we have

$$E_2 = \frac{1}{n}\left(\frac{1}{2z_0} + \frac{1}{2}\phi_f^\iota(z_0) + C_0 + C_1(b-1)\right) + O\left(\frac{\log n}{n^2}\right). \qquad \square$$

Lemma 5.23. *We have*

$$E_3 = \frac{1-b}{n} + O\left(\frac{1}{n^2}\right).$$

Proof. By explicit computations, $A_0(z) = (b-1)z^2 + O(z^3)$, so that $A_0(y_j) = (b-1)j^{-2} + O(j^{-3})$. Therefore

$$E_3 = (1-b)\sum_{j=n}^{\infty} j^{-2} + O(j^{-3})$$

and

$$\sum_{j=n}^{\infty} j^{-3} = O\left(\int_n^\infty \frac{dx}{x^3}\right) = O\left(\frac{1}{n^2}\right).$$

Similarly,

$$\sum_{j=n}^{\infty} j^{-2} \sim \int_n^\infty \frac{dx}{x^2} = \frac{1}{n},$$

so that $E_3 = (1-b)/n + O(1/n^2)$. $\square$

Finally, putting together the three preceding lemmas, the proof of Proposition 5.20 is finished. $\square$

5B2. *Outgoing part.* We will now work to obtain estimates for the outgoing part of the orbit, that is, for $n \le j \le 2n+1$. The method is largely similar to the incoming case. Recall that the estimates we obtain only depend on the chosen compact set $K \subset \mathcal{B}_f$.

We will first need a rough preliminary estimate on the boundedness of z_{2n+1}. Of course, by [Astorg et al. 2016], we know that z_{2n+1} converges to $\mathcal{L}(z_0)$, and we could deduce this preliminary estimate from there. However, we prefer to present here a direct argument, so that the proof of Theorem 5.33 remains self-contained.

Proposition 5.24. *There exists $k \in \mathbb{N}$ (independent from n) such that z_{2n+1-k} belongs to a repelling petal $\mathbb{D}(r,r)$ for f. In particular, $z_{2n+1} = O(1)$.*

Proof. Recall that by Proposition 5.20, we have that

$$\phi_n^\iota(z) := \frac{Z_n^\iota}{\epsilon_n} - \frac{1}{\epsilon_n}\sum_{j=0}^{n-1}\epsilon_j = \phi^\iota(z_0) + o(1) = O(1).$$

In particular,

$$Z_n^\iota = \left(\sum_{j=0}^{n-1}\epsilon_j\right) + O(\epsilon_n) = 1 + O\left(\frac{1}{n}\right)$$

and therefore $Z_n^o = -1 + O(1/n)$.

Let R_n denote the rectangle defined by the conditions $-1 - C/n \le \mathrm{Re}(Z) \le -3/n$ and $-1 \le \mathrm{Im}(Z) \le 1$, where $C > 0$ is a constant chosen large enough that $Z_n^o \in R_n$. Let

$$j_n := \max\{k \le 2n+1 : Z_k^o \in R_n\}. \tag{26}$$

Recall that for $j \le 2n$, we have $Z_{j+1}^o = Z_j^o + A(\epsilon_j, z_j)$, and that by Proposition 5.8, we have

$$A(\epsilon_k, z_k) = \epsilon_k A_0(z_k) + O(\epsilon_k^3, \epsilon_k^3 z_k) = O(\epsilon_k z_k^2). \tag{27}$$

Moreover, for $n \le j \le j_n$, we have

$$Z_j^o = -\frac{\pi}{2n}\cot\left(\frac{\pi}{2}Z_j^o\right) + O\left(\frac{1}{n^2}\right),$$

and therefore there exists a constant $C > 0$ such that, for all $n \le j \le j_n$,

$$|A(\epsilon_j, z_j)| \le \frac{C'}{n^3}\left|\cot\left(\frac{\pi}{2}Z_j^o\right)\right|^2 \le \frac{C}{|Z_j|^2 n^3}, \tag{28}$$

and thus

$$\left|Z_j^o - Z_n^o - \sum_{k=n}^{j-1}\epsilon_k\right| \le \frac{C}{n^3}\sum_{k=n}^{j-1}\frac{1}{|Z_k|^2}. \tag{29}$$

From (29), we can prove inductively on j that, for $n \le j \le j_n$,

$$\left|Z_j^o - Z_n^o - \sum_{k=n}^{j-1}\epsilon_k\right| = O\left(\frac{1}{n}\right)$$

and hence $j_n = 2n + O(1)$.

Let $r > 0$ be small enough that $\mathbb{D}(r, r)$ is a repelling petal for f. By the argument above and the definition of R_n, we have that $Z_{j_n}^o = O(1/n)$, so that

$$z_{2n+1-k} = -\frac{\pi}{2}\cot\left(\frac{\pi}{2}Z_{2n+1-k}^o\right) + O\left(\frac{1}{n^2}\right) = \frac{1}{k + O(1)}.$$

Therefore, we can find some k bounded independently from n such that $z_{2n+1-k} \in \mathbb{D}(r, r)$. $\square$

We now introduce approximate outgoing Fatou coordinates:

Definition 5.25. For $n \leq m \leq 2n+1$, let

$$\phi_n^o(z_m) := \frac{Z_n^o}{\epsilon_n} + \frac{1}{\epsilon_n}\sum_{j=n}^{m-1} \epsilon_j .$$

Lemma 5.26. *We have*

$$\phi_n^o(z_m) = \frac{Z_m^o}{\epsilon_n} - \frac{1}{\epsilon_n}\sum_{j=n}^{m-1} A(\epsilon_j, z_j).$$

Proof. We have

$$\sum_{j=n}^{m-1} A(\epsilon_j, z_j) = \sum_{j=n}^{m-1} Z_{j+1}^o - Z_j^o - \epsilon_j = Z_m^o - Z_n^o - \sum_{j=n}^{m-1} \epsilon_j$$

so that

$$\frac{Z_n^o}{\epsilon_n} + \frac{1}{\epsilon_n}\sum_{j=n}^{m-1} \epsilon_j = \phi_n^o(z_m) = \frac{Z_m^o}{\epsilon_n} - \frac{1}{\epsilon_n}\sum_{j=n}^{m-1} A(\epsilon_j, z_j). \qquad \square$$

Proposition 5.27. *Let $k \in \mathbb{N}$ be the integer from Proposition 5.24. Let $y_{2n+1-k} := z_{2n+1-k}$ and $y_{2n+1} = f^k(y_{2n+1-k})$. For $n \leq j \leq 2n$ we define*

$$y_j := f^{-(2n+1-j)}(y_{2n+1}),$$

where f^{-1} is the local inverse of f fixing 0: $f^{-1}(z) = z - z^2 + z^3 - bz^4 + O(z^5)$. We have

$$\sum_{j=n}^{2n} A_0(z_j) - A_0(y_j) = \frac{C_1(b-1)}{n} + O\left(\frac{\log n}{n^2}\right).$$

Proof. The proof mirrors the incoming case, so we will only sketch it and leave the details to the reader. Recall that $y_{2n+1} = O(1)$ by Proposition 5.24 and that z_{2n+1-k} belongs to a repelling petal for f for some $k \in \mathbb{N}$ independent from n, so that the $(y_j)_{n \leq j \leq 2n+1}$ are well-defined.

By a straightforward adaptation of Lemma 5.15,

$$z_j - y_j = O\left(\frac{2n+1-j}{n^2}\right)$$

for $n \leq j \leq 2n+1$. More precisely, this applies for $n \leq j \leq 2n+1-k$; but it is clear from the definition of the y_j that for $2n+1-k \leq j \leq 2n+1$ we have $z_j - y_j = O(1/n^2)$. Therefore the proof of Lemma 5.16 can be repeated to yield that

$$\sum_{j=n}^{2n} A_0(z_j) - A_0(y_j) = (b-1)\sum_{j=n}^{2n} z_j^2 - y_j^2 + O\left(\frac{\log n}{n^2}\right). \tag{30}$$

Next, we have, for $n \leq j \leq 2n$,

$$\begin{aligned} z_j = (\psi^o_{\epsilon_j})^{-1}(Z^o_j) &= (\psi^o_{\epsilon_j})^{-1}\left(Z^o_{2n+1} - \sum_{k=j}^{2n} \epsilon_k + A(\epsilon_k, z_k)\right) \\ &= -\frac{\pi}{2n}\cot\left(\frac{\pi}{2} Z^o_{2n+1} - \frac{\pi}{2}\sum_{k=j}^{2n} \epsilon_k + A(\epsilon_k, z_k)\right) + O\left(\frac{1}{n^2}\right). \end{aligned}$$

Through computations similar to those appearing in the proof of Proposition 5.19, we deduce that

$$z_j = -\frac{1}{-1/z_{2n+1} - \beta_j} + O\left(\frac{1}{n^2}\right), \tag{31}$$

with

$$\beta_j := \frac{2n}{\pi}\tan\left(\frac{\pi}{2} x_j\right) = \frac{2n}{\pi}\tan\left(\frac{\pi}{2}\sum_{k=j}^{2n} \epsilon_k + A(\epsilon_k, z_k)\right).$$

On the other hand,

$$-\frac{1}{y_j} = -\frac{1}{y_{2n+1}} - \sum_{k=j}^{2n} A_0(y_j),$$

from which it follows that

$$y_j = -\frac{1}{-1/y_{2n+1} - \gamma_j},$$

with $\gamma_j := \sum_{k=j}^{2n} A_0(y_j)$. Then, again, similar computations show that

$$z_j^2 - y_j^2 = \frac{1}{n^2}\Phi(x_j) + O\left(\frac{1}{n^2(2n+1-j)}\right),$$

and $x_j = (2n - j + O(1))/n$ for $n \leq j \leq 2n$. Therefore, we finally obtain

$$\sum_{j=n}^{2n} A_0(z_j) - A_0(y_j) = \frac{b-1}{n}\int_0^1 \Phi(x)\,dx + O\left(\frac{\log n}{n^2}\right) = \frac{C_1(b-1)}{n} + O\left(\frac{\log n}{n^2}\right). \qquad \square$$

In what follows, a slight technical complication comes from the fact that the expected endpoint of the orbit, z_{2n+1}, needs not lie in a small enough repelling petal in which ϕ^o_f is well-defined. In order to overcome this issue, we stop a few iterations short and work instead with z_{2n+1-k}.

We now come to the main proposition of this subsection:

Proposition 5.28. *We have*

$$\phi^o_n(z_{2n+1-k}) = \phi^o_f(z_{2n+1-k}) + \frac{E^o(z_{2n+1-k})}{n} + O\left(\frac{\log n}{n^2}\right),$$

where $E^o(z) = -\frac{1}{2}\phi^o_f(z) - C_0 - (C_1 - 1)(b-1)$.

Proof. We proceed similarly to the proof of Proposition 5.20. We have, for z in a small enough repelling petal,

$$\phi_f^o(z) = -\frac{1}{z} - \sum_{j=1}^{\infty} A_0(f^{-j}(z)), \tag{32}$$

where f^{-1} is the inverse branch of f fixing 0. With the same notation as in Proposition 5.27, we set

$$y_j := f^{j-(2n+1-k)}(z_{2n+1-k}).$$

We have

$$\begin{aligned}\phi_n^o(z_{2n+1-k}) - \phi_f^o(z_{2n+1-k}) &= \frac{Z_{2n+1-k}^o}{\epsilon_n} + \frac{1}{z_{2n+1-k}} + \sum_{j=n}^{2n-k} -\frac{1}{\epsilon_n} A(\epsilon_j, z_j) + A_0(y_j) + \sum_{j=-\infty}^{n-1} A_0(y_j) \\ &= E_1 + E_2 + E_3,\end{aligned} \tag{33}$$

where

$$\begin{aligned} E_1 &= \frac{Z_{2n+1-k}^o}{\epsilon_n} + \frac{1}{z_{2n+1-k}}, \\ E_2 &= \sum_{j=n}^{2n-k} -\frac{1}{\epsilon_n} A(\epsilon_j, z_j) + A_0(y_j), \\ E_3 &= \sum_{j=-\infty}^{n-1} A_0(y_j). \end{aligned} \tag{34}$$

Lemma 5.29. *We have*

$$E_1 = \frac{1}{n} \frac{1}{2z_{2n+1-k}} + O\left(\frac{1}{n^2}\right).$$

Proof of the lemma. By Proposition 5.12, we have

$$Z_{2n+1-k}^o = -\frac{\epsilon_{2n+1-k}}{z_{2n+1-k}} + O\left(\frac{1}{n^3}\right)$$

so that

$$\begin{aligned} E_1 &= \frac{1}{z_{2n+1-k}} - \frac{\epsilon_{2n+1-k}}{\epsilon_n} \frac{1}{z_{2n+1-k}} + O\left(\frac{1}{n^2}\right) \\ &= \frac{1}{z_{2n+1-k}} \left(1 - \sqrt{\frac{n^2+n+O(1)}{n^2+2n+O(1)}}\right) + O\left(\frac{1}{n^2}\right) \\ &= \frac{1}{n} \frac{1}{2z_{2n+1-k}} + O\left(\frac{1}{n^2}\right). \end{aligned}$$

□

Lemma 5.30. *We have*

$$E_2 = \frac{1}{n}\left(-\frac{1}{2z_{2n+1-k}} - \frac{1}{2}\phi_f^o(z_{2n+1-k}) - C_0 - C_1(b-1)\right) + O\left(\frac{\log n}{n^2}\right).$$

Proof of the lemma. We have

$$\begin{aligned} E_2 &= \sum_{j=n}^{2n-k} A_0(y_j) - \frac{1}{\epsilon_n} A(\epsilon_j, z_j) \\ &= \left(\sum_{j=n}^{2n-k} A_0(y_j) - \frac{\epsilon_j}{\epsilon_n} A_0(z_j)\right) - \left(\sum_{j=n}^{2n-k} C_0\epsilon_j^3 + O(z_j\epsilon_j^3)\right). \end{aligned}$$

As before, we have

$$\frac{1}{\epsilon_n}\sum_{j=n}^{2n-k} C_0\epsilon_j^3 + O(z_j\epsilon_j^3) = \frac{C_0}{n} + O\left(\frac{1}{n^2}\right).$$

On the other hand, we have

$$\sum_{j=n}^{2n-k} A_0(y_j) - \frac{\epsilon_j}{\epsilon_n} A_0(z_j) = \sum_{j=n}^{2n-k}\left(1 - \frac{\epsilon_j}{\epsilon_n}\right) A_0(z_j) + \sum_{j=n}^{2n-k} A_0(y_j) - A_0(z_j).$$

Now note that

$$\sum_{j=n}^{2n-k}\left(1 - \frac{\epsilon_j}{\epsilon_n}\right) A_0(z_j) = \sum_{j=n}^{2n-k}\left(1 - \frac{\epsilon_j}{\epsilon_n}\right) A_0(y_j) + \sum_{j=n}^{2n-k}\left(1 - \frac{\epsilon_j}{\epsilon_n}\right)(A_0(z_j) - A_0(y_j)),$$

and that

$$\left|\sum_{j=n}^{2n-k}\left(1 - \frac{\epsilon_j}{\epsilon_n}\right)(A_0(z_j) - A_0(y_j))\right| \le \max_{n\le j\le 2n-k}\left|1 - \frac{\epsilon_j}{\epsilon_n}\right| \cdot \sum_{j=n}^{2n-k} |A_0(z_j) - A_0(y_j)| = O\left(\frac{1}{n^2}\right),$$

by Proposition 5.27. Therefore, as in the proof of Proposition 5.20,

$$\begin{aligned} \sum_{j=n}^{2n-k} A_0(y_j) - \frac{\epsilon_j}{\epsilon_n} A_0(z_j) &= -\frac{C_1(b-1)}{n} + \frac{1}{2n}\sum_{j=n}^{2n-k} A_0(y_j) + O\left(\frac{\log n}{n^2}\right) \\ &= -\frac{1}{n}\left(C_1(b-1) + \frac{1}{2z_{2n+1-k}} + \frac{1}{2}\phi_f^o(z_{2n+1-k})\right) + O\left(\frac{\log n}{n^2}\right) \end{aligned}$$

from which the lemma follows. □

Lemma 5.31. *We have*

$$E_3 = \frac{b-1}{n} + O\left(\frac{1}{n^2}\right).$$

Proof of the lemma. The proof is the same as in the incoming case: it follows from the fact that $A_0(y) = (b-1)y^2 + O(y^3)$ and

$$y_j = \frac{1}{2n-j} + O\left(\frac{1}{(2n-j)^2}\right).$$ □

This completes the proof of Proposition 5.28. □

5B3. *Conclusion.*

Proposition 5.32. *We have*

$$\frac{1}{\epsilon_n}\left(\left(\sum_{j=0}^{2n}\epsilon_j\right)-2\right)=-\frac{1}{n}\left(\frac{1}{2}+\phi_g^\iota(w_0)\right)+O\left(\frac{1}{n^2}\right).$$

Proof. We have

$$\begin{aligned}\sum_{j=0}^{2n}\epsilon_j&=\sum_{j=0}^{2n}\frac{1}{\sqrt{n^2+j+\phi_g^\iota(w_0)+o(1)}}\\&=\sum_{j=0}^{2n}\frac{1}{n}\left(1-\frac{j}{2n^2}-\frac{\phi_g^\iota(w_0)}{2n^2}+o\left(\frac{1}{n^2}\right)\right)\\&=2+\frac{1}{n}-\frac{\phi_g^\iota(w_0)}{n^2}+o\left(\frac{1}{n^2}\right)-\frac{1}{2n^3}\sum_{j=0}^{2n}j\\&=2+\frac{1}{n}-\frac{\phi_g^\iota(w_0)}{n^2}+o\left(\frac{1}{n^2}\right)-\frac{1}{2n^3}\frac{2n(2n+1)}{2}\\&=2-\frac{\phi_g^\iota(w_0)}{n^2}-\frac{1}{2n^2}+o\left(\frac{1}{n^2}\right).\end{aligned}$$

On the other hand

$$\frac{1}{\epsilon_n}=\sqrt{n^2+n+O(1)}=n\left(1+\frac{1}{2n}+O\left(\frac{1}{n^2}\right)\right)=n+\frac{1}{2}+O\left(\frac{1}{n}\right),$$

and therefore

$$\begin{aligned}\frac{1}{\epsilon_n}\left(\left(\sum_{j=0}^{2n}\epsilon_j\right)-2\right)&=\left(-\frac{\phi_g^\iota(w_0)}{n^2}-\frac{1}{2n^2}+o\left(\frac{1}{n^2}\right)\right)\left(n+\frac{1}{2}+O\left(\frac{1}{n}\right)\right)\\&=-\frac{1}{n}\left(\frac{1}{2}+\phi_g^\iota(w_0)\right)+O\left(\frac{1}{n^2}\right).\end{aligned}$$

□

We are now finally ready to prove the following theorem:

Theorem 5.33 (Lavaurs' theorem with an error estimate). *Let $K\subset\mathcal{B}_f\times\mathcal{B}_g$ be a compact set. For all $(z_0,w_0)\in K$ and all sufficiently large n we have*

$$z_{2n+1}=\mathcal{L}_f(z_0)+\frac{h(z_0,w_0)}{n}+O\left(\frac{\log n}{n^2}\right),$$

where

$$h(z,w)=\frac{\mathcal{L}_f'(z)}{(\phi_f^\iota)'(z)}\left(2C_0+2(C_1-1)(b-1)-\tfrac{1}{2}+\phi_f^\iota(z)-\phi_g^\iota(w)\right)$$

is holomorphic on $\mathcal{B}_f\times\mathcal{B}_g$ and the constant in $O((\log n)/n^2)$ is independent of the point (z_0,w_0) and the integer n.

Proof. We have, by definition

$$\phi_n^o(z_{2n+1-k})=\frac{1}{\epsilon_n}Z_n^o+\frac{1}{\epsilon_n}\sum_{j=n}^{2n-k}\epsilon_j=\frac{Z_n^\iota}{\epsilon_n}-\frac{2}{\epsilon_n}+\frac{1}{\epsilon_n}\sum_{j=n}^{2n-k}\epsilon_j\;=\phi_n^\iota(z_0)-\frac{2}{\epsilon_n}+\frac{1}{\epsilon_n}\sum_{j=0}^{2n-k}\epsilon_j,$$

and therefore

$$\phi_f^o(z_{2n+1-k}) + \frac{E^o(z_{2n+1-k})}{n} = \phi_f^\iota(z_0) + \frac{E^\iota(z_0)}{n} - \frac{\phi_g^\iota(w_0) + \frac{1}{2}}{n} - \frac{1}{\epsilon_n} \sum_{j=2n-k}^{2n} \epsilon_j + O\left(\frac{\log n}{n^2}\right),$$

by Propositions 5.20, 5.28 and 5.32.

On the other hand, we have

$$\begin{aligned} \frac{1}{\epsilon_n} \sum_{j=2n-k}^{2n} \epsilon_j &= \frac{1}{1/n - 1/(2n^2) + O(1/n^3)} \left(\frac{k}{n} - k\frac{2n}{2n^3} + O\left(\frac{1}{n^3}\right)\right) \quad \text{(by Lemma 5.2)} \\ &= \left(1 + \frac{1}{2n} + O\left(\frac{1}{n^2}\right)\right)\left(k - \frac{k}{n} + O\left(\frac{1}{n^2}\right)\right) \\ &= k - \frac{k}{2n} + O\left(\frac{1}{n^2}\right). \end{aligned}$$

Therefore

$$\phi_f^o(z_{2n+1-k}) + k + \frac{E^o(z_{2n+1-k}) - \frac{1}{2}k}{n} = \phi_f^\iota(z_0) + \frac{E^\iota(z_0)}{n} - \frac{\phi_g^\iota(w_0) + \frac{1}{2}}{n} + O\left(\frac{\log n}{n^2}\right).$$

Recall that the outgoing Fatou coordinate ϕ_f^o has a well-defined inverse $\psi_f : \mathbb{C} \to \mathbb{C}$ satisfying the functional equation $\psi_f(Z+1) = f \circ \psi_f(Z)$. Observe that since $k = O(1)$, we have

$$\psi_f(\phi_f^o(z_{2n+1-k}) + k) = f^k(z_{2n+1-k}) + O\left(\frac{1}{n^2}\right) = z_{2n+1} + O\left(\frac{1}{n^2}\right).$$

Therefore, composing on both sides by ψ_f and setting $E^o(z_{2n+1}) := E^o(z_{2n+1-k}) - \frac{1}{2}k$, we get

$$\begin{aligned} z_{2n+1} &= (\phi_f^o)^{-1}\left(\phi_f^\iota(z_0) + \frac{E^\iota(z_0) - E^o(z_{2n+1}) - \frac{1}{2} - \phi_g^\iota(w_0)}{n} + O\left(\frac{\log n}{n^2}\right)\right) \\ &= \mathcal{L}_f(z_0) + ((\phi_f^o)^{-1})'(\phi_f^\iota(z_0))\left(\frac{E^\iota(z_0) - E^o(z_{2n+1}) - \frac{1}{2} - \phi_g^\iota(w_0)}{n}\right) + O\left(\frac{\log n}{n^2}\right) \\ &= \mathcal{L}_f(z_0) + \frac{\mathcal{L}_f'(z_0)}{(\phi_f^\iota)'(z_0)}\left(\frac{E^\iota(z_0) - E^o(z_{2n+1}) - \frac{1}{2} - \phi_g^\iota(w_0)}{n}\right) + O\left(\frac{\log n}{n^2}\right). \end{aligned}$$

In particular, we have proved that $z_{2n+1} = \mathcal{L}_f(z_0) + O(1/n)$. From there, we deduce that

$$\phi_f^o(z_{2n+1-k}) + k = \phi_f^o(z_0) + O\left(\frac{1}{n}\right).$$

Plugging this into the expression for $E^o(z_{2n+1})$, we finally obtain

$$z_{2n+1} = \mathcal{L}_f(z_0) + \frac{1}{n}\frac{\mathcal{L}_f'(z_0)}{(\phi_f^\iota)'(z_0)}\left(2C_0 + 2(C_1 - 1)(b-1) - \tfrac{1}{2} + \phi_f^\iota(z_0) - \phi_g^\iota(w_0)\right) + O\left(\frac{\log n}{n^2}\right). \qquad \square$$

5C. *Choice of index.* Assume that z_0 is a Siegel fixed point for the Lavaurs map $\mathcal{L}_f$, and let λ be its multiplier. Denote by κ_{z_0} the index from Theorem 2.2: it is given by the formula

$$\kappa_{z_0} = \frac{2b_2c_0}{\lambda(1-\lambda)} + \frac{c_1}{\lambda},$$

with $2b_2 = \mathcal{L}''_f(z_0),\ c_0 = h(z_0),\ c_1 = h'(z_0)$, and

$$h(z) := \frac{\mathcal{L}'_f(z)}{(\phi^\iota_f)'(z)}\big(2C_0 + 2(C_1-1)(b-1) - \tfrac{1}{2} + \phi^\iota_f(z) - \phi^\iota_g(w_0)\big). \tag{35}$$

The function h is the error term computed in the previous section.

A straightforward computation gives us that

$$\kappa_{z_0} = 1 + \frac{C + \phi^\iota_f(z_0) - \phi^\iota_g(w_0)}{((\phi^\iota_f)'(z_0))^2}\left(\frac{\mathcal{L}''_f(z_0)(\phi^\iota_f)'(z_0)}{\lambda(1-\lambda)} - (\phi^\iota_f)''(z_0)\right) \tag{36}$$

for some universal constant $C \in \mathbb{R}$.

Observe that $\mathrm{Re}(\kappa_{z_0})$ is independent from w_0 if and only if

$$\frac{\mathcal{L}''_f(z_0)(\phi^\iota_f)'(z_0)}{\lambda(1-\lambda)} - (\phi^\iota_f)''(z_0) = 0. \tag{37}$$

If condition (37) is satisfied, then $\kappa_{z_0} = 1$, and accordingly, Theorem 2.2 implies that there are no wandering domains for P converging to the bi-infinite orbit of $(z_0, 0)$, since we are then in the expulsion scenario of the trichotomy.

On the other hand, if the equality (37) is not satisfied, then $w_0 \mapsto \kappa_{z_0}(w_0)$ is a nonconstant holomorphic function (defined on the parabolic basin $\mathcal{B}_g$) of the form $w_0 \mapsto a\phi^\iota_g(w_0) + b$, with $a, b \in \mathbb{C}$ (independent from w_0) and $a \neq 0$. Therefore, the condition for $\mathrm{Re}(\kappa_{z_0}(w_0))$ to be negative is equivalent to $\phi^\iota_g(w_0)$ belonging to some half-plane, but $\phi^\iota_g(\mathcal{B}_g)$ contains a domain of the form

$$U := \{W \in \mathbb{C} : \mathrm{Re}(W) > R - k|\mathrm{Im}(W)|\}$$

for some $R > 0$ and $k \in (0, 1)$; see, e.g., [Shishikura 2000, Proposition 2.2.1, p. 330]. Since U intersects any open half-plane, if condition (37) is not satisfied, then there exists some open subset $U_0 \subset U$ for which $\mathrm{Re}(\kappa_{z_0}(w_0)) < 0$, and so by Theorem 2.2 there is a wandering domain accumulating on $(z_0, 0)$.

6. A Lavaurs map with a Siegel disk

The goal of this section is to construct a polynomial f of the form $f(z) = z + z^2 + z^3 + O(z^4)$, whose Lavaurs map has a Siegel fixed point with Diophantine multiplier λ, which does not satisfy equality (37). The outline of the argument is as follows:

- We start by finding a degree-7 real polynomial whose Lavaurs map has a superattracting fixed point, and for which a suitable reformulation of (37) does not hold.

- We perturb that polynomial to get an attracting but not superattracting fixed point, in a way that equality (37) still does not hold.
- We apply quasiconformal surgery to get a multiplier arbitrarily close to 1.
- We show that in the limit, we get a polynomial whose Lavaurs map has a parabolic fixed point that does not exit the parabolic basin.
- We perturb that last polynomial to get a Siegel fixed point, leaving the family of real polynomials; we prove that condition (37) does not hold for that last polynomial.

Recall that in [Astorg et al. 2016], there are two constructions of a Lavaurs map with an attracting fixed point. One is based on a residue computation near infinity in the Ecalle cylinder and makes use of the fact that in the family $f_a(z) := z + z^2 + az^3$, the multiplier of the horn map e_a of f_a at the ends of the Ecalle cylinder is a nonconstant holomorphic function of a. This method cannot be used in a family of polynomials of the form $f(z) = z + z^2 + z^3 + O(z^4)$, where those fixed points for the horn map are persistently parabolic. This is why we adapt the second strategy for the first two steps described above.

Remark. From now on we will be using slightly different notation than in previous sections. Namely we will drop the subscript f from the Fatou coordinates and the Lavaurs map in order to have space for other indexes in the subscript. It will be clear from the context to which function the Fatou coordinates or Lavaurs maps correspond.

Let ϕ^ι be the incoming Fatou coordinate, and ψ^o the outgoing Fatou parametrization. Recall that the Lavaurs map is given by $\mathcal{L} = \psi^o \circ \phi^\iota : \mathcal{B}_f \to \mathbb{C}$, the lifted horn map is $\mathcal{E} = \phi^\iota \circ \psi^o : V \to \mathbb{C}$, with $V \subset \mathbb{C}$ containing $\{Z : |\mathrm{Im}(Z)| > R\}$ for R large enough. We have $\mathcal{E} \circ \phi^\iota = \phi^\iota \circ \mathcal{L}$, and $\mathcal{E}(Z+1) = \mathcal{E}(Z) + 1$, so $\mathcal{E}$ descends to a self-map of $\mathbb{C}/\mathbb{Z}$. Conjugating by the isomorphism $Z \mapsto e^{2i\pi Z}$, we obtain a map $e : U - \{0, \infty\} \to \mathbb{C}^*$, where U is an open set containing 0 and ∞. The map extends to U, and fixes 0 and ∞. Since we consider polynomials with $f(z) = z + z^2 + z^3 + O(z^4)$, both of those fixed points have multiplier 1.

6A. *Construction a polynomial.* Let $f(z) = z + z^2 + O(z^3)$ be a polynomial and $\zeta = \mathcal{L}(\zeta)$ a fixed point of its Lavaurs map, with multiplier λ.

Definition 6.1. If $\lambda \notin \{0, 1\}$ we say that the pair (f, ζ) is degenerate if and only if

$$\frac{\mathcal{L}''(\zeta)(\phi^\iota)'(\zeta)}{\lambda(1-\lambda)} - (\phi^\iota)''(\zeta) = 0. \tag{38}$$

Lemma 6.2. *We have*

$$\frac{\mathcal{L}''(\zeta)(\phi^\iota)'(\zeta)}{\lambda(1-\lambda)} - (\phi^\iota)''(\zeta) = \frac{\lambda}{1-\lambda}\left[\frac{(\psi^o)''(\phi^\iota(\zeta))(\phi^\iota)'(\zeta)}{(\psi^o)'(\phi^\iota(\zeta))^2} + (\phi^\iota)''(\zeta)\right]. \tag{39}$$

Proof. Since $\mathcal{L} = \psi^o \circ \phi^\iota$ we obtain

$$\begin{aligned}
\mathcal{L}'(z) &= (\psi^o)'(\phi^\iota(z))\phi'(z),\\
\mathcal{L}''(z) &= (\psi^o)''(\phi^\iota(z))\phi'(z)^2 + (\psi^o)'(\phi^\iota(z))(\phi^\iota)''(z).
\end{aligned}$$

Recalling that $\mathcal{L}'(\zeta)=\lambda$ it follows that

$$\frac{\phi'(\zeta)}{\lambda}=\frac{1}{(\psi^o)'(\phi^\iota(\zeta))},$$

and so

$$\frac{\mathcal{L}''(\zeta)\phi'(\zeta)}{\lambda}=\frac{(\psi^o)''(\phi^\iota(\zeta))(\phi^\iota)'(\zeta)^2}{(\psi^o)'(\phi^\iota(z))}+(\phi^\iota)''(\zeta).$$

It follows that

$$\begin{aligned}\frac{\mathcal{L}''(\zeta)(\phi^\iota)'(\zeta)}{\lambda(1-\lambda)}-(\phi^\iota)''(\zeta)&=\frac{1}{1-\lambda}\left[\frac{(\psi^o)''(\phi^\iota(\zeta))(\phi^\iota)'(\zeta)^2}{(\psi^o)'(\phi^\iota(\zeta))}+(\phi^\iota)''(\zeta)\right]-(\phi^\iota)''(\zeta)\\&=\frac{(\psi^o)''(\phi^\iota(\zeta))(\phi^\iota)'(\zeta)^2}{(1-\lambda)(\psi^o)'(\phi^\iota(\zeta))}+(\phi^\iota)''(\zeta)\frac{\lambda}{1-\lambda}\\&=\frac{\lambda}{1-\lambda}\left[\frac{(\psi^o)''(\phi^\iota(\zeta))(\phi^\iota)'(\zeta)}{(\psi^o)'(\phi^\iota(\zeta))^2}+(\phi^\iota)''(\zeta)\right].\end{aligned}$$

□

For the rest of the paper we shall set

$$\mathcal{F}(f,\zeta):=\frac{(\psi^o)''(\phi^\iota(\zeta))(\phi^\iota)'(\zeta)}{(\psi^o)'(\phi^\iota(\zeta))^2}+(\phi^\iota)''(\zeta), \tag{40}$$

where ψ^o and ϕ^ι are the Fatou parametrization and coordinates associated to f. Note that for $\lambda\notin\{0,1\}$ the pair (f,ζ) is degenerate if and only if $\mathcal{F}(f,\zeta)=0$.

We record here for later use the following lemma:

Lemma 6.3. *Let $f(z)=z+z^2+az^3+O(z^4)$ and let ϕ^ι denote its incoming Fatou coordinate. Let c be a critical point in the parabolic basin of f. Then we have $(\phi^\iota)''(c)=0$ if and only if either c is multiple critical point of f, or if the orbit of c meets another critical point of f.*

Proof. The sequence of functions

$$\phi_n(z):=-\frac{1}{f^n(z)}-n-(1-a)\log n$$

converges locally uniformly on the parabolic basin to

$$\phi^\iota(z):=\lim_{n\to\infty}\phi_n(z).$$

Therefore $(\phi^\iota)''(c)$ equals $\lim_{n\to\infty}\phi_n''(c)$. Moreover, $\phi_n'(z)=(f^n)'(z)/[f^n(z)]^2$ and

$$\begin{aligned}\phi_n''(c)&=\frac{d}{dz}\Big|_{z=c}\frac{(f^n)'(z)}{[f^n(z)]^2}\\&=\frac{(f^n)''(c)[f^n(c)]^2-2[(f^n)'(c)]^2f^n(c)}{[f^n(c)]^4}\\&=\frac{(f^n)''(c)}{[f^n(c)]^2}=f''(c)\frac{\prod_{k=1}^{n-1}f'(f^k(c))}{[f^n(c)]^2}.\end{aligned}$$

For the third and fourth equalities we used the fact that $f'(c) = 0$. Since c is in the parabolic basin of f, we have $[f^n(c)]^2 \sim 1/n^2$. Moreover, for $k \geq k_0$ with k_0 large enough, $f'(f^k(c)) \neq 0$ and

$$f'(f^k(c)) = 1 - \frac{2}{k} + O\left(\frac{\log k}{k^2}\right) = \exp\left(-\frac{2}{k} + O\left(\frac{\log k}{k^2}\right)\right).$$

Therefore

$$\prod_{k=k_0}^{n-1} f'(f^k(z)) = \prod_{k=k_0}^{n-1} \exp\left(-\frac{2}{k} + O\left(\frac{\log k}{k^2}\right)\right) = \frac{\exp(O(1))}{n^2}.$$

In particular,

$$\lim_{n\to\infty} \frac{\prod_{k=k_0}^{n-1} f'(f^k(c))}{[f^n(c)]^2} \neq 0,$$

so $(\phi^\iota)''(c) = 0$ if and only if $f''(c) = 0$ or $(f^k)'(c) = 0$, which concludes the proof. □

For $t \in \mathbb{R}$, a real polynomial $P(z) = z + z^2 + z^3 + O(z^4)$ and $n > \deg P$ odd, let

$$f_t(z) = P(z) - \frac{P'(t)}{nt^{n-1}} z^n.$$

Note that $f_t'(t) = 0$; the choice of this family ensures that we have a marked critical point in $\mathbb{R}$. By $\mathcal{L}_t$ we denote the Lavaurs map of phase 0 for the polynomial f_t.

Proposition 6.4. *Assume that there exists P, n and $t_\infty < 0$ as above such that*:

(1) $f_{t_\infty}(t_\infty) = 0$.

(2) $(d/dt)|_{t=t_\infty} f_t(t) < 0$.

(3) *f_{t_∞} has negative leading coefficient.*

(4) *There exists $x > 0$ in the repelling petal of f_{t_∞} that escapes to infinity.*

Then there is a sequence $t_n \to t_\infty$ such that $\mathcal{L}_{t_n}(t_n) = t_n$.

Proof. We will rely on the following two claims:

Claim 1. *For $t \in (t_\infty, t_\infty + \epsilon)$ with $\epsilon > 0$ small enough, the critical point t is in the parabolic basin of f_t.*

Proof of the claim. It is enough to show that there is $r > 0$ such that $(-r, 0)$ is in the parabolic basin of f_t for all t close enough to t_∞. Indeed, by (1) and (2), we have that for all $r > 0$ there exists $\epsilon > 0$ such that $f_t(t) \in (-r, 0)$ for all $t \in (t_\infty, t_\infty + \epsilon)$. Let

$$r_t := \sup\{r > 0 : \text{for all } y \in (-r, 0),\ 0 < f_t(y)/y < 1\}.$$

For all $y \in (-r_t, 0)$, we have $t < f_t(y) < 0$; hence y is in the parabolic basin of f_t. Finally, $t \mapsto r_t$ is continuous and $r_{t_\infty} > 0$. □

Claim 2. *There exists a sequence $\tilde{t}_n \to t_\infty$ (with $\tilde{t}_n > t_\infty$) such that $\mathcal{L}_{\tilde{t}_n}(\tilde{t}_n) = f^n_{\tilde{t}_n}(x)$.*

Proof of the claim. We adapt here the argument from [Astorg et al. 2016]. The desired equality $\mathcal{L}_{\tilde{t}_n}(\tilde{t}_n) = f_{\tilde{t}_n}^n(x)$ is equivalent to

$$\psi_{\tilde{t}_n}^o \circ \phi_{\tilde{t}_n}^\iota(\tilde{t}_n) = \psi_{\tilde{t}_n^o}(\phi_{\tilde{t}_n}^o(x) + n).$$

In particular, it is enough to find $\tilde{t}_n$ such that $\phi_{\tilde{t}_n}^\iota(\tilde{t}_n) = \phi_{\tilde{t}_n}^o(x) + n$. We look for $\tilde{t}_n$ under the form

$$\tilde{t}_n = t_\infty - \frac{\alpha}{n+u}, \quad \text{with } \alpha = \frac{1}{\frac{d}{dc}\big|_{c=t_\infty} f_c(c)}.$$

By the preceding claim, it is in the parabolic basin for n large enough.

We have $\phi_{\tilde{t}_n}^o(x) + n = n + \phi_{t_\infty}^o(x) + o(1)$ since the map $t \mapsto \phi_t^o$ is continuous. Additionally,

$$\begin{aligned}
\phi_{\tilde{t}_n}^\iota(\tilde{t}_n) &= \phi_{\tilde{t}_n}^\iota(f_{\tilde{t}_n}(\tilde{t}_n)) - 1 \\
&= -\frac{1}{f_{\tilde{t}_n}(\tilde{t}_n)} - 1 + o(1) \quad \text{(according to the asymptotic expansion of } \phi^\iota) \\
&= n + u - 1 + o(1).
\end{aligned}$$

Therefore, we have reduced the problem to solving the equation $u - 1 + o(1) = \phi_{t_\infty}^o(x)$ for $u \in \mathbb{R}$, where the $o(1)$ term is a continuous function of u. By the intermediate value theorem there is a solution $u = u_n \in (\phi_{t_\infty}^o(x), \phi_{t_\infty}^o(x) + 2)$. We can take $\tilde{t}_n = t_\infty - \alpha/(n + u_n)$, and since $(u_n)_{n \in \mathbb{N}}$ is bounded from below, the sequence (t_n) is well-defined for n large enough and converges to t_∞. □

We now come back to the proof of Proposition 6.4. For n large enough, $G_{\tilde{t}_n}(x) > 0$ (by continuity of the Green's function G). Therefore $\mathcal{L}_{\tilde{t}_n}(\tilde{t}_n) = f_{\tilde{t}_n}^n(x)$ tends to ∞, and more precisely, $+\infty$ or $-\infty$ depending on the parity of n, thanks to condition (3). Therefore the continuous function $F(t) := \mathcal{L}_t(t) - t$ alternates sign between two consecutive $\tilde{t}_n$, so by the intermediate value theorem must have a zero t_n between them. □

Proposition 6.5. *Let $P(z) := z + z^2 + z^3 + \frac{23}{7}z^4 + \frac{17}{7}z^5$, let $t_\infty := -1$ and let $n := 7$. Then P, n and t_∞ satisfy conditions* (1)–(4) *in Proposition 6.4.*

Proof. Observe that $f_{t_\infty}(t_\infty) = 0$ and $P'(t_\infty) = 1$. That second property implies that f_{t_∞} has negative leading coefficient. Therefore, conditions (1) and (3) are satisfied.

Let us check that condition (2) is also satisfied. We have

$$\frac{d}{dt}\Big|_{t=t_\infty} f_t(t) = \frac{d}{dt}\Big|_{t=t_\infty} P(t) - \frac{t}{n}P'(t) = \frac{n-1}{n}P'(t_\infty) - \frac{t_\infty}{n}P''(t_\infty) = \tfrac{6}{7} + \tfrac{1}{7}P''(-1) = -\tfrac{50}{49} < 0.$$

Finally, condition (4) is satisfied for $x := 1$. Indeed, recall here that if $f(z) = \sum_{k=0}^n a_k z^k$ is a complex polynomial and

$$R = \max\left\{1, \frac{1 + |a_0| + \cdots + |a_{n-1}|}{|a_n|}\right\}$$

then for all $|z| > R$ we have $|f(z)| \geq |z|^n/R$; hence if an orbit at any point leaves the disk of radius R, then it must converge to infinity. Observe that for our polynomial f_{t_∞} we have $R = 68$ and that a straightforward computation yields $f_{t_\infty}(1) = \frac{60}{7}$ and $|f_{t_\infty}^2(1)| > 68$.

This proves rigorously that $x := 1$ has unbounded orbit under f_{t_∞}. □

Lemma 6.6. *For $\epsilon_0 > 0$ small enough, there exists $\boldsymbol{t} > -1$ such that the following properties hold for $f_{\boldsymbol{t}}$:*

(1) *$\mathcal{L}_{\boldsymbol{t}}$ has a fixed point $x_{\boldsymbol{t}}$ with multiplier $\epsilon_0 \neq 0$.*

(2) *$\mathcal{F}(f_{\boldsymbol{t}}, x_{\boldsymbol{t}}) \neq 0$.*

(3) *$f_{\boldsymbol{t}}$ has four real critical points, ordered from left to right, c_1, c_2, c_3, c_4, with $\boldsymbol{t} = c_2$, and two nonreal complex conjugate critical points c' and $\bar{c}'$.*

(4) *The critical points c_1 and c_4 lie in the basin of infinity; the critical points c_2 and c_3 are in the parabolic basin.*

(5) *There is a unique repelling fixed point $\xi \in (c_1, c_2)$, and the intersection of $\mathbb{R}$ and the immediate basin of attraction of 0 is $(\xi, 0)$.*

(6) *There is a unique $y \in (\xi, c_2)$ such that $f_{\boldsymbol{t}}(y) = c_2$.*

Proof. We will find $\boldsymbol{t}$ by taking a perturbation of one of the t_{n_0} constructed above, with n_0 large enough.

First, note that properties (3)–(6) hold for

$$f := f_{t_\infty} : z \mapsto z + z^2 + z^3 + \tfrac{23}{7} z^4 + \tfrac{17}{7} z^5 - \tfrac{z^7}{7};$$

we leave the details to the reader; see Figure 2. Therefore, for n_0 large enough, properties (3)–(6) still hold for $f_{t_{n_0}}$, as these properties are clearly open (in $\mathbb{R}$) near $t = t_\infty$. To lighten the notation, we let $f := f_{t_{n_0}}$ and $c_2 := t_{n_0}$.

We now claim that $\mathcal{F}$ is well-defined at (f, c_2), and that $\mathcal{F}(f, c_2) \neq 0$. According to Lemma 6.3, since f satisfies conditions (3)–(6), we have $(\phi^\iota)''(c) \neq 0$. Indeed, c_2 is a simple critical point of f, and we claim that the forward orbit of c_2 does not meet any other critical point of f. To see this, note that the critical point c_2 is simple for f, and real. Since c' and $\bar{c}'$ are not real, the orbit of c_2 cannot land on either of them. Since the critical points c_1 and c_4 do not belong to the parabolic basin, the orbit of c_2 cannot land on them either. Finally, since $f(c_2) > c_3$, and since $f(c_2)$ belongs to a small attracting petal in which the sequence of iterates $(f^n(c_2))_{n \in \mathbb{N}}$ is increasing, the orbit of c_2 cannot land on c_3 either.

Now that we have proved that $(\phi^\iota)''(c_2) \neq 0$, it is sufficient to prove that

$$\frac{(\psi^o)''(\phi^\iota(c_2))(\phi^\iota)'(c_2)}{(\psi^o)'(\phi(c_2))^2} = 0.$$

In fact, since $(\phi^\iota)'(c_2) = 0$, it is suffices to prove that $(\psi^o)'(\phi^\iota(c_2)) \neq 0$. Recall that for any $Z \in \mathbb{C}$, $(\psi^o)'(Z) = 0$ if and only if there exists $n \geq 1$ such that $(\psi^o)'(Z - n)$ is a critical point for f; here, $Z = \phi^\iota(c_2)$ and $\psi^o \circ \phi^\iota(c_2) = c_2$, so we must prove that for all $n \geq 1$ and any critical point c_i of f, $f^n(c_i) \neq c_2$. Since c_1 and c_4 escape, neither of their orbits can land on c_2, and since c_2 is not periodic under f, its own orbit cannot land on itself either. Since c_3 is in the immediate parabolic basin, the orbit $(f^n(c_3))_{n \in \mathbb{N}}$ is increasing, and so does not contain c_2 since $c_3 > c_2$.

Finally, it remains to argue that the orbits of the two nonreal critical points c' and $\bar{c}'$ do not eventually land on c_2. To see that it cannot be the case, note that since the horn map e of f has two parabolic fixed points at 0 and ∞ corresponding to the ends of the Ecalle cylinder, each of those fixed points must attract singular values of e distinct from themselves; see [Astorg et al. 2016]. The singular values of e are

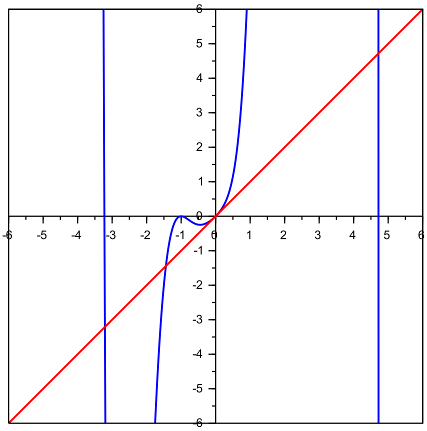

Figure 2. The graph of $f := f_{t_\infty}$ (blue), with the line $y = x$ in red. We have $c_1 \approx -2.8$, $c_2 = -1$, $c_3 \approx -0.4$, and $c_4 \approx 4$. The critical values $f(c_1)$ and $f(c_4)$ are out of the picture.

the fixed points at 0 and ∞, as well as the $\pi(c_i)$, where c_i are the critical points of f in the parabolic basin and $\pi(z) = e^{2i\pi\phi^\iota(z)}$. If $f^n(c') = c_2$ for some $n \geq 1$, then by real symmetry we would also have $f^n(\bar{c}') = c_2$, and so $\pi(c') = \pi(\bar{c}') = \pi(c_2)$; but then $\pi(c_3)$ would be the only nonfixed singular value of e, which is impossible.

Therefore f has no critical relation, and so $(\psi^o)'(\phi^\iota(c_2)) \neq 0$, and $\mathcal{F}(f, c_2) \neq 0$ as announced.

To summarize, we have proved that for n_0 large enough, the polynomial $f_{t_{n_0}}$ satisfies properties (2)–(6). Since t_{n_0} is a superattracting fixed point of $\mathcal{L}_{t_{n_0}}$ but persistently fixed, for $\epsilon_0 > 0$ small enough, there exists $\boldsymbol{t}$ close to t_{n_0} such that $f_{\boldsymbol{t}}$ satisfies (1), and by openness, if ϵ_0 is small enough, $f_{\boldsymbol{t}}$ still satisfies (2)–(6). □

The next step is to use quasiconformal deformations to construct an immersed disk D in parameter space passing through $f_{\boldsymbol{t}}$, made of polynomials p_u, whose Lavaurs map has an attracting fixed point of multiplier $e^{2i\pi u}$, $u \in \mathbb{H}$. We purposely use the notation p_u instead of f_t to emphasize the fact that except for $f_{\boldsymbol{t}}$, the polynomials p_u do not a priori belong to the family $(f_t)_{t \in \mathbb{R}^*}$.

Proposition 6.7. *Let $p := f_{\boldsymbol{t}}$ and $\epsilon_0 > 0$ be as in Lemma 6.6. There exists a holomorphic map $\Phi : \mathbb{H} \to \mathcal{P}_7$ such that*:

(1) *$\Phi(u_0) = p$ for some $u_0 \in \mathbb{H}$ with $e^{2i\pi u_0} = \epsilon_0$.*

(2) *For all $u \in \mathbb{H}$, the Lavaurs map of $\Phi(u) =: p_u$ has a fixed point z_u of multiplier $e^{2i\pi u} \in \mathbb{D}^*$, and $u \mapsto z_u$ is holomorphic.*

(3) *All the maps p_u are quasiconformally conjugated to p, the conjugacy being holomorphic outside of the grand orbit under p of the attracting basin of $z_{u_0} := x_{\boldsymbol{t}}$.*

(4) *If $e^{2i\pi u} \in (0, 1)$, then the conjugacy preserves the real line.*

(5) *The set $\Phi(\mathbb{H})$ is relatively compact in $\mathcal{P}_7$.*

Proof. Let $e : U \to \mathbb{P}^1$ be the horn map of g; since $\mathcal{L}$ has an attracting fixed point $z_{u_0} := x_{\boldsymbol{t}}$, so does e (since they are semiconjugated). Denote this attracting fixed point by x.

Let $u \in \mathbb{H}$ and μ be a Beltrami form invariant by e (i.e., $e^*\mu = \mu$) such that the corresponding quasiconformal homeomorphism h_μ conjugates e to some holomorphic map e_μ with an attracting fixed point of multiplier $e^{2i\pi u}$: $h_\mu \circ e = e_\mu \circ h_\mu$ and $e'_\mu(h_\mu(x)) = e^{2i\pi u}$. We recall here briefly how to construct such a Beltrami form, and refer the reader to [Branner and Fagella 2014] for more details. If τ is a linearizing coordinate for the horn map e near x, i.e., a holomorphic map defined near p satisfying the functional equation $\tau \circ e = \epsilon_0 \tau$, we set

$$\mu = \mu(u) := \tau^*\left(\frac{u-u_0}{u+u_0}\frac{z}{\bar z}\frac{d\bar z}{dz}\right), \tag{41}$$

where $u_0 \in \mathbb{H}$ is any point such that $e^{2i\pi u_0} = \epsilon_0$. Notice that $u \mapsto \mu(u)$ is holomorphic. In the rest of the proof, we fix $u \in \mathbb{H}$ and just use the notation μ instead of $\mu(u)$.

We choose the normalization of h_μ so that it fixes 0, 1 and ∞. Let $E(z) := e^{2i\pi z}$ and $T_1(z) := z+1$. We define

(1) $\nu := E^*\mu$ so that $\nu = T_1^*\nu$, and $\nu = \mathcal{E}^*\nu$,

(2) $\sigma := \phi^*\nu$ so that $\sigma = g^*\sigma$ and $\sigma = \mathcal{L}^*\sigma$,

(3) the quasiconformal homeomorphisms h_ν and h_σ associated to ν, σ respectively.

Since $\nu = T_1^*\nu$, the map $h_\nu \circ T_1 \circ h_\nu^{-1} : \mathbb{C} \to \mathbb{C}$ is holomorphic; since it is conjugated to T_1, it is also a translation (distinct from the identity), and we choose the normalization of h_ν so that $h_\nu \circ T_1 \circ h_\nu^{-1} = T_1$ and $h_\nu(0) = 0$. Similarly, since $\sigma = g^*\sigma$, the map $p_u := h_\sigma \circ p \circ h_\sigma^{-1}$ is holomorphic, and hence a polynomial (since it has the same topological degree as f); it also has a parabolic fixed point with one attracting petal at the origin. We choose the unique normalization of h_σ such that $p_u(z) = z + z^2 + O(z^3)$. We set $\Phi(u) := p_u$; the holomorphic dependence $u \mapsto \mu(u)$ and the parametric version of the Ahlfors–Bers theorem imply that Φ is holomorphic on $\mathbb{H}$.

We now define

(1) $\phi_\sigma := h_\nu \circ \phi \circ h_\sigma^{-1} : h_\sigma(B) \to \mathbb{C}$, where B is the parabolic basin of f,

(2) $\psi_\nu := h_\sigma \circ \psi \circ h_\nu^{-1} : \mathbb{C} \to \mathbb{C}$.

Lemma 6.8. *The map ϕ_σ is an incoming Fatou coordinate for p_u, and the map ψ_ν is an outgoing Fatou parametrization for p_u.*

Proof of the lemma. We start with ϕ_σ. First, note that since $\sigma = \phi^*\nu$, the map ϕ_σ is holomorphic on $B_\sigma := h_\sigma(B)$, which is exactly the parabolic basin of p_u. Then, note that

$$\begin{aligned}\phi_\sigma \circ p_u &= h_\nu \circ \phi \circ h_\sigma^{-1} \circ p_u = h_\nu \circ \phi \circ g \circ h_\sigma^{-1} \\ &= h_\nu \circ T_1 \circ \phi \circ h_\sigma^{-1} = T_1 \circ h_\nu \circ \phi \circ h_\sigma^{-1} = T_1 \circ \phi_\sigma.\end{aligned}$$

So ϕ_σ conjugates p_u on the whole parabolic basin to a translation, which means it is a Fatou coordinate.

The proof is completely analogous for ψ_ν: first, to prove that ψ_ν is holomorphic, note that $\nu = \psi^*\sigma$. Indeed, $\nu = \mathcal{E}^*\nu = \psi^*\phi^*\nu = \psi^*\sigma$. To conclude, one can check directly that $\psi_\nu \circ T_1 = p_u \circ \psi_\nu$. $\square$

As a consequence of the lemma, $\mathcal{E}_\nu := h_\nu \circ \mathcal{E} \circ h_\nu^{-1}$ is a lifted horn map of p_u, and $\mathcal{L}_\sigma := h_\sigma \circ \mathcal{L} \circ h_\sigma^{-1}$ is a Lavaurs map of p_u, and they have the same phase. The phase could a priori be a nonzero, but we will prove that it is not the case. In order to do that, first we will prove that $E \circ \mathcal{E}_\nu = e_\mu \circ E$, i.e., that e_μ is a horn map that lifts to $\mathcal{E}_\nu$.

Since $\nu = E^*\mu$, the map $E_\nu := h_\mu \circ E \circ h_\nu^{-1} : \mathbb{C} \to \mathbb{C}^*$ is holomorphic. Moreover, since $E : \mathbb{C} \to \mathbb{C}^*$ is a universal cover, so is E_ν. So E_ν is of the form $E_\nu(z) = \lambda e^{\alpha z}$, and with our choices of normalizations we find $E_\nu(z) = e^{2i\pi z} = E(z)$. So $E \circ h_\nu = h_\mu \circ E$.

From this, we deduce

$$\begin{aligned} E \circ \mathcal{E}_\nu &= E \circ h_\nu \circ \mathcal{E} \circ h_\nu^{-1} = h_\mu \circ E \circ \mathcal{E} \circ h_\nu^{-1} \\ &= h_\mu \circ e \circ E \circ h_\nu^{-1} = h_\mu \circ e \circ h_\mu^{-1} \circ E = e_\mu \circ E. \end{aligned}$$

Finally, it remains to observe that since e_μ is topologically conjugated to e, it also has two parabolic fixed points at 0 and ∞ respectively, each of multiplier 1. Recall that the horn map of phase 0 of a parabolic polynomial $f(z) = z + z^2 + az^3 + O(z^4)$ has multipliers at 0 and ∞ both equal to $e^{2\pi^2(1-a)}$, and that the horn map of phase $\varphi \in \mathbb{C}/\mathbb{Z}$ is obtained from the horn map e of phase 0 by multiplication by $e^{2i\pi\varphi}$. In particular, its multipliers at 0 and ∞ are respectively $e^{2\pi^2(1-a)+2i\pi\varphi}$ and $e^{2\pi^2(1-a)-2i\pi\varphi}$. In this case, since both multipliers are equal to 1, we must have $a = 1$ and $\varphi = 0$. Therefore, $\mathcal{L}_\sigma$ is the Lavaurs map of phase 0 of p_u, and $p_u(z) = z + z^2 + z^3 + O(z^4)$.

Finally, if $\pi_\sigma(z) := e^{2i\pi\phi_\sigma(z)}$, then $\pi_\sigma \circ \mathcal{L}_\sigma = e_\mu \circ \pi_\sigma$, and π_σ is locally invertible near $z_u := h_\sigma(z_{u_0})$, and $\pi_\sigma(z_u) = h_\mu(x)$. Therefore, z_u as a fixed point of $\mathcal{L}_\sigma$ has the same multiplier $e^{2i\pi u}$ as $h_\mu(x)$. This proves claims (1)–(3) of the proposition.

To prove claim (4), note that if $e^{2i\pi u} \in (0, 1)$ then the Beltrami form

$$\frac{u - u_0}{u + u_0} \frac{z}{\bar{z}} \frac{d\bar{z}}{dz}$$

has real symmetry (since then $(u - u_0)/(u + u_0) \in \mathbb{R}$). We claim that this implies that σ has real symmetry. Indeed, since $g(\mathbb{R}) = \mathbb{R}$, its Lavaurs map $\mathcal{L}$ maps a small interval $I \subset \mathbb{R}$ centered at x_t into itself. Moreover, the map $\tau \circ \pi$ semiconjugates $\mathcal{L}$ to the multiplication by $\epsilon_0 > 0$; so $\tau \circ \pi$ maps I into $\mathbb{R}$, which means that the holomorphic map $\tau \circ \pi$ is real: $\tau \circ \pi(\bar{z}) = \overline{\tau \circ \pi(z)}$ for all z in the parabolic basin of g. Therefore

$$\sigma = (\tau \circ \pi)^* \left(\frac{u - u_0}{u + u_0} \frac{z}{\bar{z}} \frac{d\bar{z}}{dz} \right)$$

has real symmetry; hence h_σ restricts to a real homeomorphism.

Finally, $\Phi : \mathbb{H} \to \mathcal{P}_7$ is bounded in the space of polynomials of degree 7. Indeed, by [Bassanelli and Berteloot 2011, Proposition 4.4] the set of polynomials of given degree with given values of the Green's function at the critical points is bounded, and since the conjugacy between the p_u and p is analytic outside of the parabolic basin, their Green's functions have the same values at critical points. □

Proposition 6.9. *With the same notation as before, there exists p_0 in the closure of $\Phi(\mathbb{H})$ such that the Lavaurs map of p_0 has a parabolic fixed point of multiplier* 1.

Proof. Applying Proposition 6.7 with $u_n = i/n$, we get a sequence of polynomials p_{u_n} such that $p_{u_n}(z) = z + z^2 + z^3 + O(z^4)$, and the Lavaurs map $\mathcal{L}_n$ of p_{u_n} has a fixed point x_n of multiplier $e^{-2\pi/n}$.

Each of the p_{u_n} are quasiconformally conjugate to the real polynomial f_t from Lemma 6.6 by a homeomorphism whose restriction to the real line is real and increasing, so the p_{u_n} still satisfy the properties (3)–(6) from Lemma 6.6.

By item (5) in Proposition 6.7, the sequence $(p_{u_n})_{n\in\mathbb{N}}$ is bounded in the space of degree-7 polynomials. So up to extracting, we may assume that:

(1) p_{u_n} converges to a degree 7 polynomial p_0.

(2) The critical points $c_{i,n}$ of p_{u_n} converge to critical points c_i of p_0.

(3) The repelling fixed point ξ_n converges to a nonattracting fixed point ξ of p_0.

(4) x_n converges to $x \in \mathbb{R}$ and y_n to $y \in \mathbb{R}$.

We denote by $\mathcal{L}$ the Lavaurs map of p_0. If we can prove that x lies in the parabolic basin of p_0, then we will get that $\mathcal{L}(x) = x$ and $\mathcal{L}'(x) = 1$. To do that, it is enough to prove that $x \in (\xi, 0)$. But for all n, we have

$$\xi_n < y_n < x_n < c_{2,n} < 0;$$

hence $\xi < y \leq x \leq c_2 < 0$. The inequality $\xi < y$ is strict because as a limit of repelling fixed points, we have $|f'(\xi)| \geq 1$, so we cannot have $y = \xi$, for otherwise we would have $\xi = f(\xi) = f(y) = c_2$ and so $f'(\xi) = 0$, a contradiction. Similarly, we cannot have $c_2 = 0$ since $p_0'(0) = 1 \neq 0$. So $x \in (\xi, 0)$ and ξ is in the parabolic basin of f, and so $\mathcal{L}'(x) = 1$ and $\mathcal{L}(x) = x$. Therefore p_0 has the desired property. ☐

Proposition 6.10. *There exists a polynomial $g(z) = z + z^2 + z^3 + O(z^4)$ of degree 7 such that*

(1) *$\mathcal{L}$ has a Siegel fixed point ζ with Diophantine multiplier, and*

(2) *the pair (g, ζ) is nondegenerate.*

Proof. Recall that $\mathcal{P}_7$ denotes the space of degree-7 polynomials of the form $f(z) = z + z^2 + z^3 + O(z^4)$, and let $V = \{(f, \zeta) \in \mathcal{P}_7 \times \mathbb{C} : \zeta \in \mathcal{B}_f\}$. V may be identified with an open set in $\mathbb{C}^5$. Finally, we consider $F := \{(f, \zeta) \in V : \mathcal{L}(\zeta) = \zeta\}$, which is an analytic hypersurface of V.

We consider the functions $\lambda : F \to \mathbb{C}$ and $\mathcal{F} : F \to \mathbb{C}$ defined as $\lambda(f, \zeta) = \mathcal{L}'(\zeta)$ and

$$\mathcal{F}(f, \zeta) = \frac{(\psi^o)''(\phi^\iota(\zeta))(\phi^\iota)'(\zeta)}{(\psi^o)'(\phi^\iota(\zeta))^2} + (\phi^\iota)''(\zeta),$$

where ϕ^ι and ψ^o are the Fatou coordinate and parametrization of f. The function λ is analytic on F, and $\mathcal{F}$ is meromorphic on F and analytic on $\lambda^{-1}(\mathbb{C}^*)$, since $(\psi^o)'(\phi^\iota(z)) = 0$ implies that $\mathcal{L}'(z) = 0$.

Let $\Phi : \mathbb{H} \to \mathcal{P}_7$ be the map defined in Proposition 6.7, and let $\widetilde{\Phi} : \mathbb{H} \to F$ be the map given by $\Phi(u) = (p_u, z_u)$, where z_u is the fixed point of the Lavaurs map of p_u with multiplier $e^{2i\pi u}$. Then $D := \widetilde{\Phi}(\mathbb{H})$ is contained in one irreducible component F_0 of F.

Let p_0 be the polynomial given by Proposition 6.9 such that its Lavaurs map has a parabolic fixed point z_0. By Proposition 6.9, (p_0, z_0) is in the closure of D in V; therefore $(p_0, z_0) \in F_0$.

Assume for a contradiction that all pairs $(f, \zeta) \in F_0$ for which $\mathcal{L}'(\zeta)$ has modulus 1 and Diophantine argument are degenerate. Then by the density of Diophantine numbers on the real line, we must have

$\mathcal{F}(f,\zeta) = 0$ on $\lambda^{-1}(S^1) \cap F_0$. Since for all $u \in \mathbb{H}$, we have $\lambda \circ \widetilde{\Phi}(u) = e^{2i\pi u}$, the analytic map λ is nonconstant on F_0. In particular, $\lambda^{-1}(S^1)$ is a real-analytic subset of F_0 of real codimension 1, nonempty since $\lambda(p_0, z_0) = 1$. By Proposition 6.7, D contains (f_t, x_t), where f_t is the polynomial given by Lemma 6.6, and such that $\mathcal{F}(f_t, z_t) \neq 0$. So the analytic map $\mathcal{F}$ is not identically zero on F_0, and therefore it cannot vanish identically on $\lambda^{-1}(S^1) \cap F_0$, a contradiction. □

References

[Astorg et al. 2016] M. Astorg, X. Buff, R. Dujardin, H. Peters, and J. Raissy, "A two-dimensional polynomial mapping with a wandering Fatou component", *Ann. of Math.* (2) **184**:1 (2016), 263–313. MR Zbl

[Bassanelli and Berteloot 2011] G. Bassanelli and F. Berteloot, "Distribution of polynomials with cycles of a given multiplier", *Nagoya Math. J.* **201** (2011), 23–43. MR Zbl

[Bedford et al. 2017] E. Bedford, J. Smillie, and T. Ueda, "Semi-parabolic bifurcations in complex dimension two", *Comm. Math. Phys.* **350**:1 (2017), 1–29. MR Zbl

[Bracci and Zaitsev 2013] F. Bracci and D. Zaitsev, "Dynamics of one-resonant biholomorphisms", *J. Eur. Math. Soc.* **15**:1 (2013), 179–200. MR Zbl

[Branner and Fagella 2014] B. Branner and N. Fagella, *Quasiconformal surgery in holomorphic dynamics*, Cambridge Stud. Adv. Math. **141**, Cambridge Univ. Press, 2014. MR Zbl

[Bruno 1971] A. D. Bruno, "Analytic form of differential equations, I", *Tr. Mosk. Mat. Obs.* **25** (1971), 119–262. In Russian; translated in *Trans. Moscow Math. Soc.* **25** (1971), 131–288. MR Zbl

[Bruno 1972] A. D. Bruno, "Analytic form of differential equations, II", *Tr. Mosk. Mat. Obs.* **26** (1972), 199–239. In Russian; translated in *Trans. Moscow Math. Soc.* **26** (1972), 199–238. MR Zbl

[Hakim 1998] M. Hakim, "Analytic transformations of $(\mathbb{C}^p, 0)$ tangent to the identity", *Duke Math. J.* **92**:2 (1998), 403–428. MR Zbl

[Ji 2020] Z. Ji, "Non-wandering Fatou components for strongly attracting polynomial skew products", *J. Geom. Anal.* **30**:1 (2020), 124–152. MR Zbl

[Jonsson 1999] M. Jonsson, "Dynamics of polynomial skew products on $\mathbb{C}^2$", *Math. Ann.* **314**:3 (1999), 403–447. MR Zbl

[Lavaurs 1989] P. Lavaurs, *Systèmes dynamiques holomorphiques*: *explosion des points périodiques*, Ph.D. thesis, Université Paris-Sud, 1989.

[Lilov 2004] K. Lilov, *Fatou theory in two dimensions*, Ph.D. thesis, University of Michigan, 2004, available at https://www.proquest.com/docview/305178400.

[López-Hernanz and Rosas 2020] L. López-Hernanz and R. Rosas, "Characteristic directions of two-dimensional biholomorphisms", *Compos. Math.* **156**:5 (2020), 869–880. MR Zbl

[Peters and Raissy 2019] H. Peters and J. Raissy, "Fatou components of elliptic polynomial skew products", *Ergodic Theory Dynam. Systems* **39**:8 (2019), 2235–2247. MR Zbl

[Peters and Smit 2018] H. Peters and I. M. Smit, "Fatou components of attracting skew-products", *J. Geom. Anal.* **28**:1 (2018), 84–110. MR Zbl

[Peters and Vivas 2016] H. Peters and L. R. Vivas, "Polynomial skew-products with wandering Fatou-disks", *Math. Z.* **283**:1-2 (2016), 349–366. MR Zbl

[Shishikura 2000] M. Shishikura, "Bifurcation of parabolic fixed points", pp. 325–363 in *The Mandelbrot set*, *theme and variations*, edited by L. Tan, Lond. Math. Soc. Lect. Note Ser. **274**, Cambridge Univ. Press, 2000. MR Zbl

[Siegel 1942] C. L. Siegel, "Iteration of analytic functions", *Ann. of Math.* (2) **43** (1942), 607–612. MR Zbl

[Sullivan 1985] D. Sullivan, "Quasiconformal homeomorphisms and dynamics, I: Solution of the Fatou–Julia problem on wandering domains", *Ann. of Math.* (2) **122**:3 (1985), 401–418. MR Zbl

[Yoccoz 1995] J.-C. Yoccoz, "Théorème de Siegel, nombres de Bruno et polynômes quadratiques", pp. 3–88 in *Petits diviseurs en dimension* 1, Astérisque **231**, Soc. Math. France, Paris, 1995. MR

Received 14 Oct 2019. Revised 23 Feb 2021. Accepted 14 Jun 2021.

MATTHIEU ASTORG: matthieu.astorg@univ-orleans.fr
Institut Denis Poisson, Collegium Sciences et Techniques, Université d'Orléans, Orléans, France

LUKA BOC THALER: luka.boc@pef.uni-lj.si
Faculty of Education, University of Ljubljana, Ljubljana, Slovenia

and

Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia

HAN PETERS: hanpeters77@gmail.com
Korteweg de Vries Institute for Mathematics, University of Amsterdam, Amsterdam, Netherlands

ANALYSIS AND PDE
Vol. 16 (2023), No. 1, pp. 89–117
DOI: 10.2140/apde.2023.16.89

GAUSSIAN ANALYTIC FUNCTIONS OF BOUNDED MEAN OSCILLATION

ALON NISHRY AND ELLIOT PAQUETTE

We consider random analytic functions given by a Taylor series with independent, centered complex Gaussian coefficients. We give a new sufficient condition for such a function to have bounded mean oscillation. Under a mild regularity assumption this condition is optimal. We give as a corollary a new bound for the norm of a random Gaussian Hankel matrix. Finally, we construct some *exceptional* Gaussian analytic functions which in particular disprove the conjecture that a random analytic function with bounded mean oscillation always has *vanishing* mean oscillation.

1. Introduction

Functions with random Fourier (or Taylor) coefficients play an important role in harmonic and complex analysis, e.g., in the proof of de Leeuw, Kahane, and Katznelson [de Leeuw et al. 1977] that Fourier coefficients of continuous functions can majorize any sequence in ℓ^2. A well-known phenomenon is that series with independent random coefficients are much “nicer” than an arbitrary function would be. For example, a theorem of [Paley and Zygmund 1930, Chapter 5, Proposition 10] (see also [Kahane 1985]) states that a Fourier series with square summable coefficients and random signs almost surely represents a *subgaussian* function on the circle.

In this paper we choose to focus on one particularly nice model of random analytic functions, the *Gaussian analytic functions* (GAFs). A GAF is given by a random Taylor series

$$G(z) = \sum_{n=0}^{\infty} a_n \xi_n z^n, \tag{1}$$

where $\{\xi_n\}_{n\ge 0}$ is a sequence of independent standard complex Gaussian random variables (i.e., with density $\frac{1}{\pi}e^{-|z|^2}$ with respect to the Lebesgue measure on the complex plane $\mathbb{C}$) and where $\{a_n\}_{n\ge 0}$ is a sequence of nonnegative constants. Many of the results we cite can be extended to more general probability distributions, and it is likely that our results can be similarly generalized, but we will not pursue this here. For recent accounts of random Taylor series, many of which focus on the distributions of their zeros, see for example [Hough et al. 2009; Nazarov and Sodin 2010]. A classical book on this and related subjects is [Kahane 1985].

We are interested in properties of the sequence $\{a_n\}$ that imply various regularity and finiteness properties of the function G represented by the series (1). One of the central spaces of analytic functions

MSC2010: primary 30B20; secondary 30H35, 47B80.
Keywords: function theory on the disc, bounded mean oscillation, Gaussian analytic functions, Bloch, probability.

is H^p, those functions F on the unit disk $\mathbb{D}$ that satisfy

$$\sup_{0<r<1} \int_0^1 |F(\mathrm{Re}(\theta))|^p \, \mathrm{d}\theta < \infty,$$

where $e(\theta) = e^{2\pi i \theta}$ for $\theta \in \mathbb{R}$ (see [Duren 1970] for background). This is a class of analytic functions whose *nontangential* boundary values on $\mathbb{T} = \{z : |z| = 1\}$ exist Lebesgue a.e. and are in $L^p(\mathbb{T})$ [Duren 1970, Theorem 2.2]. An important early effort is the aforementioned paper [Paley and Zygmund 1930], in which it was established that G is almost surely in $\bigcap_{0<p<\infty} H^p$ if and only if $\{a_n\} \in \ell^2$. One should compare this result with the well-known fact that a nonrandom analytic function belongs to H^2 if and only if the sequence of its Taylor coefficients is square summable. The related question of when G is almost surely in H^∞, the *bounded* analytic functions on the unit disk, is substantially more involved (see [Marcus and Pisier 1978]).

To fix ideas, let us make for a moment a few simplifying assumptions about the coefficients $\{a_n\}$ of the series (1). We assume $a_0 = 0$, and denote by

$$\sigma_k^2 = \sum_{n=2^k}^{2^{k+1}-1} a_n^2, \quad k \in \{0, 1, 2, \dots\},$$

the total variance of the dyadic blocks of coefficients. We say that the sequence $\{a_n\}$ (or equivalently G) is *dyadic-regular* if the sequence $\{\sigma_k\}$ is *decreasing* as $k \to \infty$. It is known (see [Kahane 1985, Chapters 7 and 8]) that if G is dyadic-regular, then G is almost surely in H^∞ if and only if

$$\sum_{k=0}^{\infty} \sigma_k < \infty, \quad \text{i.e., } \{\sigma_k\} \in \ell^1. \tag{2}$$

Moreover, if the series in (2) converges, then G is almost surely *continuous* on the closed disk $\overline{\mathbb{D}}$. Hence, a bounded random series *gains* additional regularity.

For a space S of analytic functions on the unit disk, let S_G be the set of coefficients $\{a_n\}$ for which a GAF $G \in S$ almost surely. If $S \subsetneq T$ and $S_G = T_G$, then we say that GAFs have a *regularity boost* from T to S, e.g., $C_G = H_G^\infty$. This regularity boost can be viewed as a manifestation of a general probabilistic principle: a Borel probability measure on a complete metric space tends to be concentrated on a *separable* subset of that space.[1]

Clearly there is a gap between (2) and the Paley–Zygmund condition $\{\sigma_k\} \in \ell^2$. A well-known function space that lies *strictly* between H^∞ and $\bigcap_{0<p<\infty} H^p$ is the space of analytic functions of *bounded mean oscillation* or BMOA (e.g., see [Girela 2001, Equation (5.4)]). For an interval $I \subseteq \mathbb{R}/\mathbb{Z}$ and any $f \in L^1(\mathbb{T})$, put

$$M_I(f) := \fint_I \left| f(e(\theta)) - \fint_I f \right| \mathrm{d}\theta, \quad \text{where } \fint_I f := \frac{1}{|I|} \int_I f(e(\theta)) \, \mathrm{d}\theta. \tag{3}$$

[1] Under the continuum hypothesis, by the main theorem of [Marczewski and Sikorski 1948], any Borel probability measure on a metric space with the cardinality of the continuum is supported on a separable subset.

Define the seminorm on H^1

$$\|F\|_* = \sup_{I \subseteq \mathbb{R}/\mathbb{Z}} M_I(F). \tag{4}$$

The restriction of $F \in H^1$ is necessary for F to have nontangential boundary values in L^1 on the unit disk. On the subspace of H^1 in which $F(0) = 0$, this becomes a norm. We may take BMOA to be the (closed) subspace of H^1 for which $\|\cdot\|_*$ is finite.

Fefferman and Stein [1972] show the space BMOA is the dual space of H^1 with respect to the bilinear form on analytic functions of the unit disk given by

$$(F, G) = \lim_{r \to 1} \int_0^1 F(\mathrm{Re}(\theta))\overline{G(\mathrm{Re}(\theta))}\, \mathrm{d}\theta,$$

and in many aspects it serves as a convenient "replacement" for the space H^∞. However, BMOA is not separable (see [Girela 2001, Corollary 5.4]).

One of our main results is the following.

Theorem 1.1. *A dyadic-regular Gaussian analytic function G that satisfies the Paley–Zygmund condition $\{\sigma_k\} \in \ell^2$ almost surely belongs to* VMOA, *the space of analytic functions of* ***vanishing*** *mean oscillation.*

The space VMOA is the closure of polynomials (or continuous functions) in the norm $\|\cdot\|_*$, and hence it is separable. It can alternatively be characterized as the subspace of H^1 for which $\lim_{|I| \to 0} M_I^1(F) = 0$. In fact, we show that a dyadic-regular GAF with square-summable coefficients almost surely belongs to a subspace of VMOA, which we attribute to Sledd [1981].

1A. *The Sledd Space* SL.

Sledd [1981] introduced a function space, which is contained in BMOA and is much more amenable to analysis. Define the seminorm for $F \subset H^1$

$$\|F\|_{S(T)}^2 = \sup_{|x|=1} \sum_{n=0}^{\infty} |T_n \star F(x)|^2, \tag{5}$$

where $\star$ denotes convolution on $\mathbb{T}$ and $\{T_n\}$ is a certain sequence of compactly supported bump functions in Fourier space, so that $\hat{T}_n = 1$ for modes from $[2^n, 2^{n+1}]$ (see (15) for the explicit definition of $\{T_n\}$). We let SL denote the subspace of H^1 with finite $\|\cdot\|_{S(T)}$ norm; [Sledd 1981] showed that SL $\subsetneq$ BMOA.[2] Sledd proved the following result.

Theorem I [Sledd 1981, Theorem 3.2]. *If $\{\sqrt{k}\sigma_k\} \in \ell^2$, then $G \in$* VMOA *almost surely.*

Remark 1.2. Sledd proved the result for series with random signs, but his method works also in our setting. In fact his theorem shows that G is almost surely in VMOA $\cap$ SL.

We extend the analysis of the $\|\cdot\|_{S(T)}$ seminorm, and in particular find a better sufficient condition for the finiteness of $\|G\|_{S(T)}$.

[2]The function $I_F = \sum_{n=0}^{\infty} |T_n \star F(x)|^2$ is essentially what appears in Littlewood–Paley theory. For each $\frac{2}{3} < p < \infty$, finiteness of the p-norm of I_F is equivalent to being in H^p; see [Stein 1966, Theorem 5]. Thus, in some sense SL could be viewed as a natural point in the hierarchy of H^p spaces.

Theorem 1.3. *If* $\sum_{k=1}^{\infty} \sup_{n \geq k}\{\sigma_n^2\} < \infty$, *then* $G \in \mathrm{SL}$ *almost surely.*

In particular, if G is dyadic-regular and $\{\sigma_k\} \in \ell^2$, then $G \in \mathrm{SL}$. The latter condition is necessary for G to have well-defined boundary values, and so we see that under the monotonicity assumption, a GAF G which has boundary values in L^2 is in BMOA . We also note that the condition in Theorem 1.3 is strictly weaker than the one in Theorem I (see Lemma 4.9).

The Sledd space SL is nonseparable (see Proposition 3.3). The proof of Theorem I is based on a stronger condition than $\|G\|_{S(T)} < \infty$, that in addition implies that a function is in the space $\mathrm{SL} \cap \mathrm{VMOA}$.[3] We show that this is unnecessary, as a GAF which is in SL has a regularity boost.

Theorem 1.4. *If* $G \in \mathrm{SL}$ *almost surely*, *then* $G \in \mathrm{VMOA}$ *almost surely.*

Theorems 1.4 and 1.3 imply Theorem 1.1.

This could raise suspicion that there is also a regularity boost from BMOA to VMOA, which is perhaps the most natural separable subspace of BMOA. Indeed, [Sledd 1981] asks whether it is possible to construct a non-VMOA random analytic function in BMOA.

1B. *Exceptional Gaussian analytic functions.* Sledd [1981, Theorem 3.5] gives a construction of a random analytic function with square summable coefficients which is not in BMOA, and moreover is not *Bloch* (this construction can be easily adapted to GAFs). The Bloch space, $\mathcal{B}$, contains all analytic functions F on the unit disk for which

$$\|F\|_{\mathcal{B}} := \sup_{|z| \leq 1} ((1-|z|^2)|F'(z)|) < \infty. \tag{6}$$

See [Anderson et al. 1974; Girela 2001] for more background on this space. Gao [2000] provides a complete characterization of which sequences of coefficients $\{a_n\}$ give GAFs in $\mathcal{B}$.

The space $\mathcal{B}$ is nonseparable, suggesting that GAFs in $\mathcal{B}$ could concentrate on a much smaller space. Finding this space is a natural open question and does not seem obvious from the characterization in [Gao 2000]. It is known that $\mathrm{BMOA} \subset \mathcal{B}$ (see, e.g., [Girela 2001, Corollary 5.2]), and, a priori, it could be that GAFs which are in $H^2 \cap \mathcal{B}$ are automatically in BMOA. However, our following result disproves this, and also answers the aforementioned question of Sledd.

Theorem 1.5. *We have*

$$\mathrm{SL}_G \subsetneq \mathrm{VMOA}_G \subsetneq \mathrm{BMOA}_G \subsetneq (H^2 \cap \mathcal{B})_G. \tag{7}$$

Remark 1.6. From Theorem 1.3 and standard results on boundedness of Gaussian processes, we may add that $H_G^{\infty} \subsetneq \mathrm{SL}_G$. From the example in [Sledd 1981], it also follows that $(H^2 \cap \mathcal{B})_G \subsetneq H_G^2$.

We leave open the question of the existence of a natural separable subspace S of BMOA such that $\mathrm{BMOA}_G = S_G$.

[3] Specifically, [Sledd 1981] shows that under the condition in Theorem I, $\sum_{n=0}^{\infty} \sup_{|x|=1} |T_n \star F(x)|^2$ is finite, which implies $F \in \mathrm{SL} \cap \mathrm{VMOA}$.

1C. ***Some previously known results.*** Billard [1963] (see also [Kahane 1985, Chapter 5]) proved that a random analytic function with independent symmetric coefficients is almost surely in H^∞ if and only if it almost surely extends continuously to the closed unit disk.

A complete characterization of Gaussian analytic functions which are bounded on the unit disk was found by Marcus and Pisier [1978] in terms of rearrangements of the covariance function (see also [Kahane 1985, Chapter 15]). Moreover, they show the answer is the same for Steinhaus and Rademacher random series (where the common law of all $\{\xi_n\}$ is taken uniform on the unit circle and on $\{\pm 1\}$, respectively). Their criterion can be seen to be equivalent to the finiteness of Dudley's entropy integral for the process of boundary values of G on the unit circle.

The best existing sufficient conditions that we know for the sequence $\{a_n\}$ to belong to BMOA_G are due to [Sledd 1981]. The more recent paper of [Wulan 1994] treats a more general problem, which in the particular case of VMOA gives another proof of Theorem I.

1D. ***Norms of random Hankel matrices.*** A *Hankel* matrix A is any $n \times n$ matrix with the structure $A_{ij} = (c_{i+j-2})$ for some sequence $\{c_k\}_0^\infty$. The function $\phi(z) = \sum_{k=0}^\infty c_k z^{k+1}$ is referred to as the *symbol* of A. We will consider the case that $n \in \mathbb{N}$, and we will also consider the infinite case. We denote by B the Hankel operator with the same symbol on ℓ^2, which may well be unbounded. Then by a combination of results of Fefferman and Nehari (see [Peller 2003, Chapter 1] and [Holland and Walsh 1986, Part III]), there is an absolute constant M such that

$$\frac{1}{M}\|\phi\|_* \leq \|B\| \leq M\|\phi\|_*, \tag{8}$$

with $\|B\|$ the operator norm of B.

If we take $c_m = a_{m+1}\xi_{m+1}$ for all $m \geq 0$ with $\{\xi_m\}$ i.i.d. $N_{\mathbb{C}}(0, 1)$ and with $a_m \geq 0$ for all m, then ϕ is exactly the GAF G. Moreover, by combining Theorem 3.1, Remark 3.7 and Lemma 4.8, we have that there is an absolute constant $C > 0$ such that

$$\mathbb{E}\|\phi\|_*^2 \leq C \sum_{k=1}^{\infty} \sup_{m\geq k}\{\sigma_m^2\}.$$

Note that for any $n \times n$ Hankel matrix A with symbol $\phi(z) = \sum_{k=0}^\infty c_k z^{k+1}$, if B is the infinite Hankel operator with *finite* symbol $\phi_n(z) = \sum_{k=0}^{2n} c_k z^{k+1}$, then $\|A\| \leq \|B\|$ as A is the $n \times n$ upper-left corner of B. Hence, using (8),

$$\|A\| \leq \|B\| \leq M\|\phi_n\|_*,$$

and we arrive at the following corollary.

Theorem 1.7. *There is an absolute constant $C > 0$ such that if A is an $n \times n$ Hankel matrix with symbol G (see (1)) and L is the smallest integer greater than or equal to $\log_2(2n)$, then*

$$\mathbb{E}\|A\|^2 \leq C \sum_{k=0}^{L} \sup_{k\leq m\leq L} \sigma_m^2.$$

We emphasize that by virtue of (8) the problem of estimating the norm of a random Gaussian Hankel matrix is essentially equivalent to the problem of estimating the $\|\cdot\|_*$ norm of a random Gaussian polynomial.

This is particularly relevant as random Hankel and Toeplitz matrices[4] have appeared many times in the literature and have numerous applications to various statistical problems. See the discussion in [Bryc et al. 2006] for details. The particular case of Hankel matrices with symbol $G = \sum_{k=0}^{\infty} \operatorname{Re}(\xi_k) z^{k+1}$, i.e., Hankel matrices with i.i.d. Gaussian antidiagonals, is particularly well studied. In that case, [Meckes 2007] and [Nekrutkin 2013] give proofs that $\mathbb{E}\|A\| \le c\sqrt{n \log n}$. Finer results for the symmetric Toeplitz case are available in [Sen and Virág 2013].

Furthermore, Meckes [2007] gives a matching lower bound, and his method can be applied to show that (deterministically)

$$\|A\| \ge \sup_{|z|=1} \left| \sum_{j=0}^{2(n-1)} \left(1 - \frac{|n-1-j|}{n}\right) a_j \xi_j z^j \right|.$$

Fernique's theorem [Kahane 1985, Chapter 15, Theorem 5] can then be used to show that Theorem 1.7 is sharp up to multiplicative numerical constant, at least when $a_j = j^{-\alpha}$ for $\alpha \in \mathbb{R}$.

Some results for more general random symbols exist; in particular, [Adamczak 2010, Theorem 4] shows that in the setting of Theorem 1.7,

$$\mathbb{E}\|A\|^2 \le C(\log n) \sum_{m=0}^{L} \sigma_m^2, \tag{9}$$

which is always larger than the bound in Theorem 1.7; in the case that σ_m^2 is monotonically decreasing and summable, (9) differs substantially from the condition in Theorem 1.7. Note that in Theorem 1.7, the entries of the $n \times n$ Hankel matrix are independent standard *complex* Gaussian random variables, whereas [Adamczak 2010, Theorem 4] holds for non-Gaussian symbols as well.

Organization. In Section 2, we give some background theory for working with GAFs and random series. In Section 3, we give some further properties of the space SL and we give some equivalent characterizations for $G \in \mathrm{SL}$. We also prove Theorem 1.4. In Section 4, we give a sufficient condition for G to be in SL; in particular, we prove Theorem 1.3. Finally, in Section 5 we construct exceptional GAFs, and we show the inclusions in (7) are strict.

Notation. We use the expression *numerical constant* and *absolute constant* to refer to fixed real numbers without dependence on any parameters. We make use of the notation $\lesssim$ and $\gtrsim$ and $\asymp$. In particular, we say that $f(a, b, c, \dots) \lesssim g(a, b, c, \dots)$ if there is an absolute constant $C > 0$ such that $f(a, b, c, \dots) \le C g(a, b, c, \dots)$ for all $a, b, c, \dots$. We use $f \asymp g$ to mean $f \lesssim g$ and $f \gtrsim g$.

[4] A Toeplitz matrix A has the form $A_{ij} = w_{i-j}$ for some $(w_k)_{-\infty}^{\infty}$. The symbol for such a matrix is again $\sum w_k z^k$. By reordering the rows, it can be seen that a Toeplitz matrix with symbol $\sum_{-n}^{n} w_k z^k$ has the same norm as the Hankel matrix with symbol $\sum_0^{2n} w_{k-n} z^k$.

2. Preliminaries

Some of our proofs will rely on the so-called contraction principle.

Proposition 2.1 (contraction principle). *For any finite sequence (x_i) in a topological vector space V, any continuous convex $F : V \to [0, \infty]$, any i.i.d., symmetrically distributed random variables (ϵ_i), and any (α_i) real numbers in $[-1, 1]$:*

(i) $\mathbb{E} F\big(\sum_i \alpha_i \epsilon_i x_i\big) \le \mathbb{E} F\big(\sum_i \epsilon_i x_i\big)$.

(ii) *If F is a seminorm, then $\mathbb{P}\big[F\big(\sum_i \alpha_i \epsilon_i x_i\big) \ge t\big] \le 2\mathbb{P}\big[F\big(\sum_i \epsilon_i x_i\big) \ge t\big]$ for all $t > 0$.*

This is essentially [Ledoux and Talagrand 1991, Theorem 4.4], although we have changed the formulation slightly. For convenience we sketch the proof.

Proof. The mapping

$$(\alpha_1, \alpha_2, \dots, \alpha_N) \mapsto \mathbb{E} F\bigg(\sum_i \alpha_i \epsilon_i x_i\bigg)$$

is convex. Therefore it attains its maximum on $[-1, 1]^N$ at an extreme point, i.e., an element of $\{\pm 1\}^N$. By the symmetry of the distributions, for all such extreme points, the value of the expectation is $\mathbb{E} F\big(\sum_i \epsilon_i x_i\big)$, which completes the proof of the first part.

For the second part, we may without loss of generality assume that $\alpha_1 \ge \alpha_2 \ge \cdots \ge \alpha_N \ge \alpha_{N+1} = 0$ by relabeling the variables and using the symmetry of the distributions of $\{\epsilon_i\}$. Letting $S_n = \sum_{k=1}^n \epsilon_i x_i$ for any $1 \le n \le N$, we can use summation by parts to express

$$\sum_i \alpha_i \epsilon_i x_i = \sum_i \alpha_i (S_i - S_{i-1}) = \sum_i (\alpha_i - \alpha_{i+1}) S_i.$$

Hence, as F is a seminorm,

$$F\bigg(\sum_i \alpha_i \epsilon_i x_i\bigg) \le \alpha_1 \max_{1 \le i \le N} F(S_i) \le \max_{1 \le i \le N} F(S_i).$$

Using the reflection principle, it now follows that for any $t \ge 0$,

$$\mathbb{P}\Big[\max_{1 \le i \le N} F(S_i) \ge t\Big] \le 2\mathbb{P}[F(S_N) \ge t],$$

which completes the proof (see [Ledoux and Talagrand 1991, Theorem 4.4] for details). $\square$

We also need the following standard Gaussian concentration inequality.

Proposition 2.2. *Suppose that $X = (X_j)_1^n$ are i.i.d. standard complex Gaussian variables, and suppose $F : \mathbb{C}^n \to \mathbb{R}$ is a 1-Lipschitz function with respect to the Euclidean metric. Then $\mathbb{E}|F(X)| < \infty$ and, for all $t \ge 0$,*

$$\mathbb{P}[F(X) - \mathbb{E} F(X) > t] \le e^{-t^2}.$$

Proof. This follows from the real case (see [Ledoux and Talagrand 1991, (1.5)]). The real and imaginary Gaussian random variables have variance $\frac{1}{2}$, for which reason the exponent is e^{-t^2}. $\square$

Approximation of seminorms. Let $\|\cdot\|$ be a densely defined seminorm on H^2 which dominates the H^2 norm. We will say that $\|\cdot\|$ is *approximable* if there exists a sequence of polynomials $\{p_n\}$ with $\sup_{n,j} \|z^j \star p_n(z)\| \le 1$ such that for all $F \in H^2$,

$$\sup_n \|F \star p_n\| < \infty \iff \|F\| < \infty \qquad \text{and} \qquad \sup_n \|F \star p_n\| = 0 \iff \|F\| = 0. \tag{10}$$

Let V be the quotient space of $\{F \in H^2 : \|F\| < \infty\}$ by the space $\{F \in H^2 : \|F\| = 0\}$. Then both $\|\cdot\|$ and $\sup_n \|\cdot \star p_n\|$ make V into Banach spaces with equivalent topologies, by the hypotheses. Hence (10) is equivalent to

$$\text{there exists } C > 0 \text{ such that } \frac{1}{C} \sup_n \|F \star p_n\| \le \|F\| \le C \sup_n \|F \star p_n\| \text{ for all } F \in H^2, \tag{11}$$

as the inclusion map from one of these Banach spaces to the other is continuous and hence bounded.

Remark 2.3. While approximable seminorms could be formulated in greater generality, we work in the H^2 setting to appeal to general concentration of measure theory.

We say that G is an H^2-GAF if $\{a_k\} \in \ell^2$.

Proposition 2.4. *Let G be an H^2-GAF. Let $\|\cdot\|$ be any approximable seminorm. Then the following are equivalent*:

(i) $\|G\| < \infty$ *a.s.*

(ii) $\mathbb{E}\|G\| < \infty$.

(iii) $\mathbb{E}\|G\|^2 < \infty$.

Remark 2.5. We remark that these equivalences hold in great generality for a Gaussian measure in a *separable* Banach space, due to a theorem of Fernique [Ledoux 1996, Theorem 4.1]. As the spaces BMOA and $\mathcal{B}$ are not separable, we instead will appeal to this notion of approximable.

Remark 2.6. A priori it is not clear that a seminorm being finite is a measurable event with respect to the product σ-algebra generated by the Taylor coefficients of G. However, for an approximable seminorm, measurability is implied by the equivalence in (10), since $\sup_n \|G \star p_n\|$ is clearly measurable; cf. [Kahane 1985, Chapter 5, Proposition 1].

Proof. The implications (iii) $\implies$ (ii) $\implies$ (i) are trivial, and so it only remains to show that (i) $\implies$ (iii).

Let $\{p_m\}$ be the polynomials making $\|\cdot\|$ approximable. Define

$$G_m = G \star p_m, \quad \text{where } G(z) = \sum_{k=0}^{\infty} a_k \xi_k z^k.$$

Without loss of generality, we may assume that $\|a\|_{\ell_2}^2 = \sum_{k=0}^{\infty} a_k^2 = 1$. For any $m \in \mathbb{N}$, let $k_m = \deg(p_m)$, and define the function on $\mathbb{C}^{k_m}$

$$F_m(x) = F_m(x_0, x_1, \dots, x_{k_m}) = \left\| \left(\sum_{j=0}^{k_m} a_j x_j z^j \right) \star p_m(z) \right\|.$$

Then for any complex vectors $x = (x_j)_0^{k_m}$ and $y = (y_j)_0^{k_m}$, by changing coordinates one at a time and using $\sup_{n,j} \|z^j \star p_n(z)\| \le 1$,

$$|F_m(x) - F_m(y)| \le \sum_{j=0}^{k_m} a_j |x_j - y_j| \le \|a\|_{\ell_2} \|x - y\|_{\ell_2}.$$

For any $\ell \in \mathbb{N}$, the function $\max_{1 \le m \le \ell} F_m(x)$ is again 1-Lipschitz. So define for any $\ell \in \mathbb{N}$

$$H_\ell := \max_{1 \le m \le \ell} \|G_m\|.$$

Therefore, by Proposition 2.2, we have that, for all $t \ge 0$ and all $\ell \in \mathbb{N}$,

$$\mathbb{P}[|H_\ell - \mathbb{E} H_\ell| \ge t] = \mathbb{P}\big[\big|F_m(\xi_0, \xi_1, \ldots, \xi_{k_m}) - \mathbb{E}(\max_{1 \le m \le \ell} F_m(\xi_0, \xi_1, \ldots, \xi_{k_m}))\big| \ge t\big] \le 2e^{-t^2}. \tag{12}$$

Hence there is an absolute constant $C > 0$ such that for all $\ell \in \mathbb{N}$,

$$|\mathrm{med}(H_\ell) - \mathbb{E}(H_\ell)| \le C, \tag{13}$$

where $\mathrm{med}(X)$ denotes any median of the random variable X.

Suppose that $\|G\| < \infty$ a.s. By (10), $\sup_m \|G_m\| = \sup_\ell H_\ell < \infty$ a.s. Therefore there is a constant $M > 0$ such that $\mathbb{P}(\sup_\ell H_\ell > M) < \frac{1}{2}$, and so $\mathrm{med}(H_\ell) \le M$ for all $\ell \in \mathbb{N}$. By monotone convergence and (13),

$$\mathbb{E} \sup_m \|G_m\| = \mathbb{E} \sup_\ell H_\ell = \sup_\ell \mathbb{E} H_\ell \le M + C.$$

Using (11), there is another absolute constant C such that

$$\mathbb{E}\|G\| \le C\, \mathbb{E} \sup_m \|G_m\| < \infty.$$

Using (11) and (12), $\mathrm{Var}(\sup_m \|G_m\|) < \infty$, and therefore

$$\mathbb{E}\|G\|^2 \le C\, \mathbb{E}(\sup_m \|G_m\|^2) \le \mathrm{Var}(\sup_m \|G_m\|) + (\mathbb{E} \sup_m \|G_m\|)^2 < \infty. \qquad \square$$

Both $\|\cdot\|_*$ and $\|\cdot\|_{\mathcal{B}}$ are approximable with $\{p_n\}$ given by the analytic part of the Fejér kernel

$$K_n^A(z) = \sum_{k=0}^{n} \left(1 - \frac{k}{n+1}\right) z^k.$$

See [Holland and Walsh 1986, Theorems 1 and 4]. In fact, it is elementary to observe the following.

Lemma 2.7. *For any* $f \in \mathrm{H}^1(\mathbb{T})$, $\sup_n \|K_n^A \star f\|_* = \|f\|_*$ *and* $\sup_n \|K_n^A \star f\|_{\mathcal{B}} = \|f\|_{\mathcal{B}}$.

Proof. We show the first of these claimed identities. For any fixed interval $I \subseteq \mathbb{R}/\mathbb{Z}$,

$$\lim_{n \to \infty} M_I(f \star K_n^A) = M_I(f),$$

and hence, $\sup_n \|K_n^A \star f\|_* \ge \|f\|_*$. On the other hand, for any fixed $\omega \in \mathbb{T}$, $f_\omega := z \mapsto f(\bar{\omega} z)$ has that $\|f_\omega\|_* = \|f\|_*$. Hence by comparing to the Fejér kernel (see (14)), which is positive, for any $n \ge 0$,

$$\|f \star K_n^A\|_* = \|f \star K_n\|_* = \left\| \int f_{e(\theta)} K_n(e(\theta)) d\theta \right\|_* \le \sup_{\omega \in \mathbb{T}} \|f_\omega\|_* = \|f\|_*,$$

where the inequality follows as $\|\cdot\|_*$ is convex and the Fejér kernel $K_n(z)$ is the density of a probability measure on $\mathbb{T}$. ☐

Corollary 2.8. *Let F be an H^2-GAF. Then $\|F\|_* < \infty$ a.s. if and only if $\mathbb{E}\|F\|_* < \infty$, and $\|F\|_{\mathcal{B}} < \infty$ a.s. if and only if $\mathbb{E}\|F\|_{\mathcal{B}} < \infty$.*

We also have that the probability that a GAF is in BMOA, VMOA, or $\mathcal{B}$ is either 0 or 1.

Proposition 2.9. *For any H^2-GAF G, the events $\{G \in \mathrm{BMOA}\}$, $\{G \in \mathrm{VMOA}\}$, $\{G \in \mathcal{B}\}$ all have probability* 0 *or* 1.

Proof. Take the decomposition $G = G_{\le n} + G_{>n}$, where $G_{\le n}$ is the n-th Taylor polynomial of G at 0. Then as $G_{\le n}$ is a polynomial, $\|G_{\le n}\|_* < \infty$ almost surely. Hence $\|G\|_* < \infty$ if and only if $\|G_{>n}\|_* < \infty$, up to null events. Therefore, $\|G\|_* < \infty$ differs from a tail event of $\{\xi_n : 1 \le n < \infty\}$ by a null event, and so the statement follows from the Kolmogorov 0-1 law. The same proof shows that $\mathbb{P}[G \in \mathcal{B}] \in \{0, 1\}$.

For VMOA, as $G_{\le n}$ is a polynomial,

$$\lim_{|I|\to 0} \sup_I M_I^1(G_{\le n}) = 0 \quad \text{a.s.},$$

and the same reasoning as above gives the 0-1 law. ☐

3. The Sledd space

Let K_n for $n \in \mathbb{N}$ be the n-th Fejér kernel, which for $|z| = 1$ is given by

$$K_n(z) = \sum_{k=-n}^{n} \left(1 - \frac{|k|}{n+1}\right) z^k = \frac{1}{n+1} \cdot \frac{|1 - z^{n+1}|^2}{|1 - z|^2}. \tag{14}$$

This kernel has the two familiar properties: $\|K_n\|_1 = 1$ and $K_n(z) \le 4/(n+1) \cdot (1/|1-z|^2)$.

For a function $F : \mathbb{T} \to \mathbb{C}$ with a Laurent expansion on $\mathbb{T}$, let $\hat{F} : \mathbb{Z} \to \mathbb{C}$ be its Fourier coefficients, i.e., let $\hat{F}(k)$ be the k-th coefficient of its Laurent expansion.

We let T_n be the dyadic trapezoidal kernel

$$\begin{aligned} T_0(z) &= 1 + \tfrac{1}{2} z + \tfrac{1}{2} z^{-1} \\ T_n &= 2K_{2^{n+2}} - K_{2^{n+1}} + K_{2^{n-1}} - 2K_{2^n}, \quad n \ge 1. \end{aligned} \tag{15}$$

The kernel T_n satisfies that $\hat{T}_n$ is supported in $[2^{n-1}, 2^{n+2})$, has $|\hat{T}_n(K)| \le 1$ everywhere, has $\hat{T}_n(K) = 1$ for $K \in [2^n, 2^{n+1}]$, and satisfies

$$\sum_{n=0}^{\infty} \hat{T}_n(K) = 1$$

for all integers $K \ge 0$. Further, $\|T_n\|_1 \le 6$ for all $n \ge 0$. Also,

$$|T_n(z)| \le 20 \cdot 2^{-n} |1 - z|^{-2}. \tag{16}$$

Recall that in terms of the kernels $\{T_n\}$, we defined the seminorm (in (5)) as

$$\|F\|_{S(T)}^2 = \sup_{|x|=1} \sum_{n=0}^{\infty} |T_n \star F(x)|^2. \tag{17}$$

In [Sledd 1981], it is shown that this norm is related to $\|\cdot\|_*$ in the following way.

Theorem 3.1. *If $F \in H^1$, then there is an absolute constant $C > 1$ such that*

$$\|F\|_* \leq C\|F\|_{S(T)}.$$

Sledd also gives a sufficient condition for F to be in VMOA, though we observe that there is a stronger one that follows directly from Theorem 3.1.

Theorem 3.2. *If $F \in H^1$ and if*

$$\lim_{k\to\infty} \sup_{|x|=1} \sum_{n=k}^{\infty} |T_n \star F(x)|^2 = 0,$$

then $F \in$ VMOA.

Proof. The space VMOA is the closure of continuous functions in the BMOA norm. Hence it suffices to find, for any $\epsilon > 0$, a decomposition $G = G_1 + G_2$ with G_1 continuous and $\|G_2\|_{\mathrm{BMOA}} \leq \epsilon$. For any $\epsilon > 0$, we may by hypothesis pick k sufficiently large such that

$$\sup_{|x|=1} \sum_{n=k}^{\infty} |T_n \star G(x)|^2 \leq \epsilon.$$

Using Theorem 3.1, it follows that if we take the decomposition

$$G = G_1 + G_2, \qquad \text{where} \quad G_1 = \sum_{n=0}^{k-1} T_n \star G \quad \text{and} \quad G_2 = \sum_{n=k}^{\infty} T_n \star G,$$

then G_1 is a polynomial and is in particular continuous. From the properties of the Fourier support of $\{T_n\}$,

$$T_n \star G_2 = \begin{cases} T_n \star G & \text{if } n \geq k+2, \\ \sum_{p=k}^{k+3} T_n \star T_p \star G & \text{if } k-2 \leq n \leq k+1, \\ 0 & \text{if } n \leq k-3. \end{cases} \tag{18}$$

Thus we have, for any $n \in [k-2, k+1]$ by using $\|T_n\|_1 \leq 6$ and convexity of the square, that

$$\|T_n \star G_2\|_\infty^2 \lesssim \sup_{n\geq k} \|T_n \star G\|_\infty^2 \leq \sup_{|x|=1} \sum_{n=k}^{\infty} |T_n \star G(x)|^2 \leq \epsilon.$$

Applying Theorem 3.1 to G_2 and using the properties derived in (18),

$$\|G_2\|_*^2 \lesssim \sup_{|x|=1} \sum_{n=0}^{\infty} |T_n \star G_2(x)|^2 = \sup_{|x|=1} \sum_{n=k}^{\infty} |T_n \star G_2(x)|^2 \leq \sup_{|x|=1} \sum_{n=k+2}^{\infty} |T_n \star G(x)|^2 + \sum_{n=k-2}^{k+1} \|T_n \star G_2\|_\infty^2 \lesssim \epsilon. \quad \square$$

Proposition 3.3. *The Sledd space* SL *is nonseparable.*

Sketch of the proof. We sketch the construction of an uncountable family of analytic functions in SL whose pairwise distances in $\|\cdot\|_{S(T)}$ are uniformly bounded below. Put

$$G_j(z) = \frac{1}{2^j+1} z^{2^{j+1}} K_{2^j}(ze(1/j)), \quad j \geq 1.$$

Notice that $\hat{G}_j$ is supported in $[2^j, 2^{j+2}]$ and that G_j has the following properties:

(1) $|G_j(e(-1/j))| = 1$.

(2) $|G_j(e(\theta))| \leq 1$ for all θ.

(3) $|G_j(e(-1/j+\theta))| \lesssim 2^{-j}$ when $c2^{-j/2} \leq |\theta| \leq \pi$.

For any $A \subset 5\mathbb{N}$ let $H_A = \sum_{n\in A} G_n$. By the above properties all these functions belong to SL and are uniformly separated from each other. □

Remark 3.4. The construction above gives an example of functions in SL which are not continuous on the boundary of the disk.

GAFs and the Sledd space. We shall be interested in applying Sledd's condition to GAFs, for which purpose it is possible to make some simplifications. For any $n \geq 0$, let R_n be the kernel defined by

$$\hat{R}_n(K) = \begin{cases} 1 & \text{if } K \in [2^n, 2^{n+1}), \\ 0 & \text{otherwise.} \end{cases}$$

In short, for a GAF, (and more generally any random series with symmetric independent coefficients) we may replace the trapezoidal kernel T_n by R_n; specifically:

Theorem 3.5. *Suppose G is an H^2-GAF. Then the following are equivalent*:

(i) $\lim_{k\to\infty} \sup_{|x|=1} \sum_{n=k}^{\infty} |T_n \star G(x)|^2 = 0$ *a.s.*

(ii) $\lim_{k\to\infty} \mathbb{E}\big[\sup_{|x|=1} \sum_{n=k}^{\infty} |T_n \star G(x)|^2\big] = 0$.

(iii) $\lim_{k\to\infty} \mathbb{E}\big[\sup_{|x|=1} \sum_{n=k}^{\infty} |R_n \star G(x)|^2\big] = 0$.

(iv) $\lim_{k\to\infty} \sup_{|x|=1} \sum_{n=k}^{\infty} |R_n \star G(x)|^2 = 0$ *a.s.*

Proof of Theorem 3.5. We begin with the equivalence of (ii) and (iii), and the implication that (iii) implies (ii). For any $n \geq 0$ and any $j \in \{1,2,3,4\}$ define $R_{n,j} = T_n \star R_{n+j-1}$. Then $T_n = \sum_{j=1}^{4} R_{n,j}$. Using convexity, we can bound

$$\sup_{|x|=1} \sum_{n=k}^{\infty} |T_n \star G(x)|^2 \lesssim \sum_{j=1}^{4} \sup_{|x|=1} \sum_{n=k}^{\infty} |R_{n,j} \star G(x)|^2.$$

Since $\hat{R}_{n,j}$ is supported in $[2^n, 2^{n+1})$ and has $\|\hat{R}_{n,j}\|_\infty \leq 1$, the contraction principle implies that, for any $0 \leq k \leq m < \infty$,

$$\mathbb{E} \sup_{|x|=1} \sum_{n=k}^{m} |R_{n,j} \star G(x)|^2 \leq \mathbb{E} \sup_{|x|=1} \sum_{n=k}^{m} |R_n \star G(x)|^2 \leq \mathbb{E} \sup_{|x|=1} \sum_{n=k}^{\infty} |R_n \star G(x)|^2.$$

Sending $m \to \infty$ and using monotone convergence implies that

$$\mathbb{E} \sup_{|x|=1} \sum_{n=k}^{\infty} |R_{n,j} \star G(x)|^2 \le \mathbb{E} \sup_{|x|=1} \sum_{n=k}^{\infty} |R_n \star G(x)|^2,$$

from which the desired convergence follows.

Conversely, to see that (ii) implies (iii), we begin by bounding

$$\sup_{|x|=1} \sum_{n=k}^{\infty} |R_n \star G(x)|^2 \le \sum_{j=1}^{4} \sup_{|x|=1} \sum_{\substack{n \ge k \\ n \in 4\mathbb{N}+j}}^{\infty} |R_n \star G(x)|^2.$$

Then by the contraction principle and monotone convergence, for any $j \in \{1, 2, 3, 4\}$,

$$\mathbb{E} \sup_{|x|=1} \sum_{\substack{n \ge k \\ n \in 4\mathbb{N}+j}}^{\infty} |R_n \star G(x)|^2 \le \mathbb{E} \sup_{|x|=1} \sum_{\substack{n \ge k \\ n \in 4\mathbb{N}+j}}^{\infty} |T_n \star G(x)|^2 \le \mathbb{E} \sup_{|x|=1} \sum_{n \ge k}^{\infty} |T_n \star G(x)|^2,$$

which completes the proof of the desired implication.

We turn to showing the equivalence of (i) and (ii). From Markov's inequality, (ii) implies that

$$\sup_{|x|=1} \sum_{n=k}^{\infty} |T_n \star G(x)|^2 \xrightarrow[k \to \infty]{\mathbb{P}} 0.$$

As the sequence $\sup_{|x|=1} \sum_{n=k}^{\infty} |T_n \star G(x)|^2$ is monotone and therefore always converges, it follows that it converges almost surely to 0.

Define for each $k \in \mathbb{N}$ the seminorms

$$\| \cdot \|_{S(R),k} : H^1 \to [0, \infty], \quad \text{where } \|f\|_{S(R),k}^2 := \sup_{|x|=1} \sum_{n=k}^{\infty} |R_n \star f(x)|^2,$$

$$\| \cdot \|_{S(T),k} : H^1 \to [0, \infty], \quad \text{where } \|f\|_{S(T),k}^2 := \sup_{|x|=1} \sum_{n=k}^{\infty} |T_n \star f(x)|^2.$$

In preparation to use Proposition 2.4, we make the following claim.

Claim 3.6. *The seminorms* $\{\| \cdot \|_{S(R),k}, \| \cdot \|_{S(T),k}\}$ *are approximable.*

We shall return to the proof of this claim after completing the proof of Theorem 3.5. We now show the equivalence of (iii) and (iv). The proof of the equivalence of (i) and (ii) is the same. From (iii) it follows from Markov's inequality that

$$\sup_{|x|=1} \sum_{n=k}^{\infty} |R_n \star G(x)|^2 \xrightarrow[k \to \infty]{\mathbb{P}} 0.$$

By monotonicity $\sup_{|x|=1} \sum_{n=k}^{\infty} |R_n \star G(x)|^2$ converges almost surely, and so it converges almost surely to 0. From (iv) and by Claim 3.6, there exists a k_0 such that

$$\mathbb{E} \sup_{|x|=1} \sum_{n=k_0}^{\infty} |R_n \star G(x)|^2 < \infty.$$

As a consequence, it is possible to take $k_0 = 0$. By dominated convergence,

$$\lim_{k\to\infty} \mathbb{E}\left[\sup_{|x|=1} \sum_{n=k}^{\infty} |R_n \star G(x)|^2\right] = 0. \qquad \square$$

Proof of Claim 3.6. Let p_m be the polynomial of degree $2^{m+1}-1$ whose nonzero coefficients are all 1. Then, for any $m > k$,

$$\|p_m \star f\|^2_{S(R),k} = \sup_{|x|=1} \sum_{n=k}^{m} |R_n \star f(x)|^2 \xrightarrow[m\to\infty]{} \|f\|^2_{S(R),k}.$$

Let $q_m(z)$ be the sum of the analytic part of $\sum_{k=0}^{m} T_k(z)$. Then, for analytic f in the disk,

$$q_m \star f = \sum_{k=0}^{m} T_k \star f.$$

Moreover, using (15) the sum $\sum_{k=0}^{m} T_k$ can be represented by a sum of a finite number of Fejér kernels with cardinality bounded independent of m. Therefore there is an absolute constant $C > 0$ such that, for all m,

$$\|q_m \star f\|_\infty \le \left\|\sum_{k=0}^{m} T_k\right\|_1 \|f\|_\infty \le C\|f\|_\infty. \tag{19}$$

Using that $\hat{q}_m(j) = 1$, for $0 \le j \le 2^m - 1$,

$$\|q_m \star f\|^2_{S(T),k} \ge \sup_{|x|=1} \sum_{n=k}^{m-2} |T_n \star f(x)|^2 \xrightarrow[m\to\infty]{} \|f\|^2_{S(T),k},$$

and so if $\sup_m \|q_m \star f\|^2_{S(T),k} < \infty$, this means $\|f\|^2_{S(T),k} < \infty$ also. Conversely, if $\|f\|^2_{S(T),k} < \infty$, then $\sup_{n\ge k} \|T_n \star f\|_\infty < \infty$, and hence, with the same C as in (19),

$$\max_{m-1\le n\le m+2} \|q_m \star T_n \star f\|_\infty \le C\|f\|_{S(T),k}.$$

So

$$\|q_m \star f\|^2_{S(T),k} \le \sup_{|x|=1} \sum_{n=k}^{m-2} |T_n \star f(x)|^2 + \sum_{n=m-1}^{m+2} \|q_m \star T_n \star f\|^2_\infty \le (1+4C^2)\|f\|^2_{S(T),k} < \infty. \qquad \square$$

Remark 3.7. In reviewing the proof of Theorem 3.5, one also sees that under the same assumptions the following are equivalent:

(i) $\sup_{|x|=1} \sum_{n=0}^{\infty} |T_n \star G(x)|^2 < \infty$ a.s.

(ii) $\mathbb{E}\left[\sup_{|x|=1} \sum_{n=0}^{\infty} |T_n \star G(x)|^2\right] < \infty$.

(iii) $\mathbb{E}\left[\sup_{|x|=1} \sum_{n=0}^{\infty} |R_n \star G(x)|^2\right] < \infty$.

(iv) $\sup_{|x|=1} \sum_{n=0}^{\infty} |R_n \star G(x)|^2 < \infty$ a.s.

Moreover, the proof gives that there is an absolute constant $C > 0$ such that

$$\frac{1}{C}\mathbb{E}\|G\|^2_{S(R)} \le \mathbb{E}\|G\|^2_{S(T)} \le C\,\mathbb{E}\|G\|^2_{S(R)}.$$

Finally, we show that for a GAF, finiteness of $\|G\|_{S(R)}$ in fact implies $G \in \mathrm{VMOA}$.

Theorem 3.8. *If G is an H^2-GAF for which*

$$\|G\|_{S(R)}^2 = \sup_{|x|=1} \sum_{n=0}^{\infty} |R_n \star G(x)|^2 < \infty \quad a.s.,$$

then

$$\lim_{k\to\infty} \sup_{|x|=1} \sum_{n=k}^{\infty} |R_n \star G(x)|^2 = 0 \quad a.s.$$

Furthermore, $\|G\|_{S(R)} < \infty$ implies G is in VMOA.

We will need the following result [Kahane 1985, Chapter 5, Proposition 12].

Proposition 3.9. *Let $u_1, u_2, \dots$ be a sequence of continuous functions on the unit circle such that* $\limsup_{k\to\infty} \|u_k\|_\infty > 0$, *and let $\theta_1, \theta_2, \dots$ be a sequence of independent random variables uniformly distributed on* $[0, 1]$. *Then almost surely there exists a $t \in \mathbb{R}/\mathbb{Z}$ such that* $\limsup_{k\to\infty} |u_k(e(t-\theta_k))| > 0$.

Proof of Theorem 3.8. Let $v_n := |R_n \star G|^2$ for all $n \geq 1$. Suppose to the contrary that

$$V := \lim_{k\to\infty} \sup_{|x|=1} \sum_{n=k}^{\infty} v_n(x)$$

is not almost surely 0. Then as V is tail-measurable, there is a $\delta \in (0, 1)$ so that $V > \delta$ a.s. By monotonicity, it follows that, for all k,

$$\sup_{|x|=1} \sum_{n=k}^{\infty} v_n(x) > \delta \quad \text{a.s.}$$

Furthermore, deterministically,

$$\lim_{m\to\infty} \sup_{|x|=1} \sum_{n=k}^{m} v_n(x) = \sup_{|x|=1} \sum_{n=k}^{\infty} v_n(x).$$

By continuity of measure,

$$\lim_{m\to\infty} \mathbb{P}\left(\sup_{|x|=1} \sum_{n=k}^{m} v_n(x) > \delta \right) = \mathbb{P}\left(\lim_{m\to\infty} \sup_{|x|=1} \sum_{n=k}^{m} v_n(x) > \delta \right) = 1.$$

Thus there is a sequence $m_1 < m_1' < m_2 < m_2' < \cdots$ such that if $u_k := \sum_{n=m_k}^{m_k'} v_n$, then

$$\mathbb{P}(\|u_k\|_\infty > \delta) > \delta.$$

By Borel–Cantelli,

$$\mathbb{P}\left(\limsup_{k\to\infty} \|u_k\|_\infty > \delta\right) = 1.$$

Let θ_k be i.i.d. uniform variables on $[0, 1]$ which are also independent of G. Therefore by conditioning on G and using Proposition 3.9 there is almost surely a $t \in \mathbb{R}/\mathbb{Z}$ such that

$$\limsup_{k\to\infty} v_k(e(t-\theta_k)) > 0.$$

Because $\{v_n(xe(\theta_k))\}$ has the same distribution as $\{v_n(x)\}$, it follows there is almost surely a $s \in \mathbb{R}/\mathbb{Z}$ such that

$$\limsup_{k\to\infty} v_k(e(s)) > 0.$$

Therefore $\|G\|^2_{S(R)} \geq V = \infty$ a.s., which concludes the first part of the proof.

Using Theorem 3.2, Theorem 3.5 and Remark 3.7, the second conclusion follows. □

4. Sufficient condition for a GAF to be Sledd

In this section we will give a sufficient condition on the coefficients of the GAF to be in SL . Recall that a standard complex Gaussian random variable is one with density on $\mathbb{C}$ given by $\frac{1}{\pi}e^{-|z|^2}$. A vector (H_1, H_2) is a centered complex Gaussian vector if it has the same distribution as a linear image of the i.i.d. standard complex Gaussian random variables $(\xi_j : j \in \mathbb{N})$, or equivalently if it is the linear image of some pair of independent standard complex Gaussian random variables (Z_1, Z_2). We begin with the following preliminary calculation.

Lemma 4.1. *Let (H_1, H_2) be a centered complex Gaussian vector with $\mathbb{E}|H_1|^2 = \mathbb{E}|H_2|^2 = 1$ and $|\mathbb{E}[H_1\overline{H}_2]| = \rho \in [0, 1]$. Then for all $|\lambda| < (1-\rho^2)^{-1/2}$,*

$$\mathbb{E}e^{\lambda(|H_1|^2-|H_2|^2)} = \frac{1}{1-\lambda^2(1-\rho^2)}.$$

Proof. We may assume without loss of generality that $\mathbb{E}[H_1\overline{H}_2] = \rho \geq 0$. Hence, we may write

$$\begin{pmatrix} H_1 \\ H_2 \end{pmatrix} = A\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} := \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix},$$

where $Z = (Z_1, Z_2)$ are independent standard complex normals, considered as a column vector. Therefore,

$$|H_1|^2 - |H_2|^2 = Z^*A^*\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}AZ.$$

It follows that

$$\mathbb{E}e^{\lambda(|H_1|^2-|H_2|^2)} = \frac{1}{\det\big(\mathrm{Id} - \lambda A^*\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)A\big)} = \frac{1}{1-\lambda^2(1-\rho^2)}.$$ □

We shall apply this equality to the complex Gaussian process $Q_n(\theta) := R_n \star G(e(\theta))$. Then

$$\sigma_n^2 = \mathbb{E}|Q_n|^2 \quad \text{and define} \quad \rho_n := \rho_n(\theta_1-\theta_2) := \sigma_n^{-2}|\mathbb{E}[Q_n(\theta_1)\,\overline{Q_n(\theta_2)}]| \in [0, 1].$$

In the case that $\sigma_n^2 = 0$, we may take any value in $[0, 1]$ for ρ_n. From Lemma 4.1, we have, for any $|\lambda|^2 < (1-\rho_n^2)^{-1}\sigma_n^{-4}$,

$$\mathbb{E}\exp(\lambda(|Q_n(\theta_1)|^2 - |Q_n(\theta_2)|^2)) = \frac{1}{1-\lambda^2(1-\rho_n^2)\sigma_n^4}. \tag{20}$$

While we would like to use $\sigma_n^4(1-\rho_n^2(\theta_1-\theta_2))$ as a distance, it does not obviously satisfy the triangle inequality, for which reason we introduce

$$\Delta_n(\theta) := \mathbb{E}\big||Q_n(\theta)|^2 - |Q_n(0)|^2\big|, \tag{21}$$

which defines a pseudometric on $\mathbb{R}/\mathbb{Z}$ through $\Delta_n(\theta_1,\theta_2) := \Delta_n(\theta_1-\theta_2)$. While Δ_n may not obviously control the tails of $|Q_n(\theta)|^2$, we observe the following lemma.

Lemma 4.2. *There is a numerical constant $C > 1$ such that, for all choices of $\{a_k\}$ and any $n \geq 0$ and all $\theta \in [0,1]$,*

$$\frac{1}{C}\sigma_n^2\sqrt{1-\rho_n^2(\theta)} \leq \Delta_n(\theta) \leq C\sigma_n^2\sqrt{1-\rho_n^2(\theta)}.$$

Proof. From (20), it follows that

$$\mathbb{E}(|Q_n(\theta)|^2 - |Q_n(0)|^2)^2 = 2\sigma_n^4(1-\rho_n^2),$$
$$\mathbb{E}(|Q_n(\theta)|^2 - |Q_n(0)|^2)^4 = 24\sigma_n^8(1-\rho_n^2)^2.$$

Hence by Cauchy–Schwarz,

$$\Delta_n^2(\theta) \leq 2\sigma_n^4(1-\rho_n^2).$$

On the other hand, by the Paley–Zygmund inequality,

$$(|Q_n(\theta)|^2 - |Q_n(0)|^2)^2 \geq \sigma_n^4(1-\rho_n^2)$$

with probability at least $\frac{1}{4}\cdot\frac{2}{24}$ which gives a lower bound for Δ_n of the same order. □

We now define two pseudometrics on $[0,1]$ in terms of $\{\Delta_n\}$:

$$\begin{aligned} d_\infty(\theta_1,\theta_2) &:= d_\infty(\theta_1-\theta_2) := \sup_{n\geq 0}\Delta_n(\theta_1-\theta_2),\\ d_2^2(\theta_1,\theta_2) &:= d_2^2(\theta_1-\theta_2) := \sum_{n\geq 0}\Delta_n^2(\theta_1-\theta_2). \end{aligned} \tag{22}$$

Using Lemma 4.1, we can also now give a tail bound for differences of

$$F(\theta) := \sum_{n=0}^{\infty}|Q_n(\theta)|^2. \tag{23}$$

Lemma 4.3. *Let $\theta_1,\theta_2 \in [0,1]$. There is a numerical constant $C > 1$ such that, for all $t \geq 0$,*

$$\mathbb{P}[F(\theta_1) - F(\theta_2) \geq t] \leq \exp\left(-C\min\left\{\frac{t}{d_\infty(\theta_1-\theta_2)}, \frac{t^2}{d_2^2(\theta_1-\theta_2)}\right\}\right).$$

Proof. The desired tail bound follows from estimating the Laplace transform of $F(\theta_1)-F(\theta_2)$. Specifically we use the following estimate.

Lemma 4.4. *Suppose that there are $\lambda_0, \sigma > 0$ and X a real-valued random variable for which*

$$\mathbb{E}e^{\lambda X} \leq e^{\lambda^2\sigma^2/2} \quad \text{for } \lambda^2 \leq \lambda_0^2.$$

Then, for all $t \geq 0$,

$$\mathbb{P}[X \geq t] \leq \exp\left(-\min\left\{\frac{\lambda_0 t}{2}, \frac{t^2}{2\sigma^2}\right\}\right).$$

Proof. Applying Markov's inequality, for any $t \geq 0$ and $0 < \lambda \leq \lambda_0$,

$$\mathbb{P}[X \geq t] \leq \exp(-\lambda t + \lambda^2\sigma^2/2).$$

Taking $\lambda = t/\sigma^2$, if possible, gives one of the bounds. Otherwise, for $\lambda_0 \leq t/\sigma^2$, taking $\lambda = \lambda_0$ gives the other bound. □

We return to estimating the Laplace transform of $F(\theta_1) - F(\theta_2)$. Recalling (20), for any

$$|\lambda|^2 < \lambda_\star^2 := \inf_{n\in\mathbb{N}} (1-\rho_n^2)^{-1}\sigma_n^{-4} \leq \frac{C^2}{d_\infty(\theta_1-\theta_2)^2},$$

where C is the numerical constant from Lemma 4.2, we have

$$\mathbb{E}\exp(\lambda(F(\theta_1) - F(\theta_2))) = \prod_{n=1}^{\infty} \frac{1}{1-\lambda^2(1-\rho_n^2)\sigma_n^4}. \tag{24}$$

Therefore, for all $|\lambda|^2 < \lambda_\star^2/2$,

$$\mathbb{E}\exp(\lambda(F(\theta_1) - F(\theta_2))) \leq \prod_{n=1}^{\infty} \frac{1}{1-\lambda^2(1-\rho_n^2)\sigma_n^4} \leq \exp\left(2\lambda^2 \sum_{n=1}^{\infty}(1-\rho_n^2)\sigma_n^4\right), \tag{25}$$

using $(1-x)^{-1} \leq e^{2x}$ for $0 \leq x \leq \frac{1}{2}$. The desired conclusion now follows from Lemmas 4.2 and 4.4. □

We now recall the technique of Talagrand for controlling the supremum of processes. We let $T = [0, 1]$. Define, for any metric d on T and any $\alpha \geq 1$,

$$\gamma_\alpha(d) = \inf \sup_{t\in T} \sum_{k\geq 0} d(t, C_k) 2^{k/\alpha}, \tag{26}$$

where the infimum is taken over all choices of finite subsets $(C_k)_{k\geq 0}$ of T with cardinality $|C_k| = 2^{2^k}$ for $k \geq 1$ and $|C_0| = 1$.

Theorem 4.5 (see [Talagrand 2001, Theorem 1.3]). *Let d_∞ and d_2 be two pseudometrics on T and let $(X_t)_{t\in T}$ be a process so that*

$$\mathbb{P}[|X_s - X_t| \geq u] \leq 2\exp\left(-\min\left\{\frac{u}{d_\infty(s,t)}, \frac{u^2}{d_2^2(s,t)}\right\}\right).$$

Then there is a universal constant $C > 0$ such that

$$\mathbb{E}\sup_{s,t\in T} |X_s - X_t| \leq C(\gamma_1(d_\infty) + \gamma_2(d_2)).$$

Hence, as an immediate corollary of this theorem and of Lemma 4.3, we have:

Corollary 4.6. *There is a numerical constant $C > 0$ such that*

$$\mathbb{E}\sup_\theta F(\theta) \leq C(\gamma_1(d_\infty) + \gamma_2(d_2)) + \sqrt{\textstyle\sum a_n^2}.$$

Finally, we give some estimates on the quantities γ_1 and γ_2 for the metrics we consider. We begin with an elementary observation that shows $\Delta_n(\theta)$ must decay for sufficiently small angles (when $|\theta| \leq 2^{-n}$).

Lemma 4.7. *There is a numerical constant $C > 1$ such that, for all $\theta \in [-1, 1]$,*

$$1 - \rho_n^2(\theta) \le C 2^{2n} |\theta|^2 \quad \textit{and} \quad \Delta_n(\theta) \le C\sigma_n^2 2^n |\theta|.$$

Proof. We begin by observing that ρ_n can always be bounded by

$$\rho_n \ge \sigma_n^{-2} \sum_{k=2^n}^{2^{n+1}-1} |a_k|^2 \cos(2\pi k(\theta)) \ge 1 - 2\pi^2 2^{2n+2}\theta^2.$$

The proof now follows from Lemma 4.2. □

We now show that $\mathbb{E}\|G\|_{S(R)}^2$ has the desired control. For any $k \ge 0$, let

$$\tau_k^2 = \sup_{n \ge k} \sigma_n^2.$$

Lemma 4.8. *There is an absolute constant $C > 0$ such that*

$$\mathbb{E}\|G\|_{S(R)}^2 \le C \sum \tau_k^2.$$

This lemma proves Theorem 1.3.

Proof. From Corollary 4.6,

$$\mathbb{E}\|G\|_{S(R)}^2 \lesssim \gamma_1(d_\infty) + \gamma_2(d_2) + \sum \tau_k^2.$$

We will choose C_k to be the dyadic net $\{\ell 2^{-2^k} : 1 \le \ell \le 2^{2^k}\}$. Then using Lemma 4.7 it follows that for any $t \in [0, 1]$,

$$\begin{aligned} d_\infty(t, C_k) = d_\infty(2^{-2^k}) &\lesssim \sup_{n \ge 0}\{\Delta_n(2^{-2^k})\sigma_n^2\} \lesssim \sup_{n \ge 0}\{2^{-(n-2^k)_-}\sigma_n^2\}, \\ d_2^2(t, C_k) = d_2^2(2^{-2^k}) &\lesssim \sum_{n=0}^{\infty} \Delta_n^2(2^{-2^k})\sigma_n^4 \lesssim \sum_{n=0}^{\infty}\{2^{-2(n-2^k)_-}\sigma_n^4\}. \end{aligned} \tag{27}$$

In the previous equations, $x_- := -\min\{x, 0\}$.

This leads to the following estimates on γ_1 and γ_2:

$$\gamma_1(d_1) \le \sum_{k=0}^{\infty} \Big\{\sup_{n \ge 0}\{2^{-(n-2^k)_-}\sigma_n^2\}\Big\} \cdot 2^k, \tag{28}$$

$$\gamma_2(d_2) \le \sum_{k=0}^{\infty} \sqrt{\left\{\sum_{n \ge 0} 2^{-2(n-2^k)_-}\sigma_n^4\right\}} \cdot 2^{k/2}. \tag{29}$$

We show that

$$\gamma_1(d_1) + \gamma_2(d_2) \lesssim \sum_{k=0}^{\infty} \tau_k^2. \tag{30}$$

To control $\gamma_1(d_1)$, we begin by applying Cauchy condensation:

$$\gamma_1(d_1) \lesssim \sum_{k=0}^{\infty} \Big\{\sup_{n \ge 0}\{2^{-(n-k)_-}\sigma_n^2\}\Big\}. \tag{31}$$

We then estimate

$$\sup_{n\geq 0}\{2^{-(n-k)_-}\sigma_n^2\} \leq \sum_{n=0}^{k} 2^{n-k}\sigma_n^2 + \tau_k^2.$$

Applying this bound and changing the order of summation for the first, it follows that $\gamma_1(d_1) \lesssim \sum_k \tau_k^2$.

To control $\gamma_2(d_2)$, we again begin by applying Cauchy condensation which results in

$$\gamma_2(d_2) \leq \sum_{k=0}^{\infty} \sqrt{\left\{\sum_{n\geq 0} 2^{-2(n-k)_-}\sigma_n^4\right\} \cdot \frac{1}{k}}. \tag{32}$$

We then split the sum to get

$$\gamma_2(d_2) \leq \sum_{k=0}^{\infty} \sqrt{\left\{\sum_{0\leq n\leq k} 2^{2(n-k)}\tau_n^4\right\} \cdot \frac{1}{k}} + \sum_{k=0}^{\infty} \sqrt{\left\{\sum_{n\geq k} \tau_n^4\right\} \cdot \frac{1}{k}}. \tag{33}$$

To the first term we apply the subadditivity of $\sqrt{\cdot}$, which produces

$$\sum_{k=0}^{\infty} \sqrt{\left\{\sum_{0\leq n\leq k} 2^{2(n-k)}\tau_n^4\right\} \cdot \frac{1}{k}} \lesssim \sum_{k=0}^{\infty} \frac{1}{\sqrt{k}} \cdot \left\{\sum_{0\leq n\leq k} 2^{n-k}\tau_n^2\right\} \lesssim \sum_{n=0}^{\infty} \frac{1}{\sqrt{n}} \cdot \tau_n^2,$$

where the second sum follows from changing the order of summation. To the second term in (33) we again apply Cauchy condensation:

$$\sum_{k=0}^{\infty} \sqrt{\sum_{n\geq k} \tau_n^4 \cdot \frac{1}{k}} \lesssim \sum_{k=0}^{\infty} \sqrt{\left\{\sum_{j\geq k} \tau_{2^j}^4 \cdot 2^j\right\}} \cdot 2^{k/2} \lesssim \sum_{k=0}^{\infty} \left\{\sum_{j\geq k} \tau_{2^j}^2 \cdot 2^{j/2}\right\} \cdot 2^{k/2} \lesssim \sum_{j=0}^{\infty} \tau_{2^j}^2 \cdot 2^j,$$

where the penultimate inequality follows from subadditivity of $\sqrt{\cdot}$ and the final inequality follows by changing the order of summation. From another application of Cauchy condensation, (30) follows. $\square$

We remark that sequences for which $\sum_{k=0}^{\infty} \tau_k^2 = \infty$ but which are square summable necessarily have some amount of lacunary behavior.

Lemma 4.9. *Suppose $\sum_{k=0}^{\infty} \tau_k^2 = \infty$ but $\sum_{n=0}^{\infty} \sigma_n^2 < \infty$. Then for any $C > 1$ there is a sequence $\{j_k\}$ tending to infinity with $j_{k+1}/j_k > C$ for all k such that*

$$\sum_{k=1}^{\infty} \sigma_{j_k}^2 \cdot j_k = \infty.$$

Proof. Using Cauchy condensation, we have that, for any $m \in \mathbb{N}$ with $m > 1$,

$$\sum_{j=1}^{\infty} \tau_{m^j}^2 \cdot m^j = \infty = \sum_{k=0}^{\infty} \tau_k^2.$$

Let $\{j_k^*\}$ be the subsequence of $\{m^j\}$ at which $\tau_{m^j} > \tau_{m^{j+1}}$. Picking j_k as an ℓ in $[j_k^*, mj_k^*)$ that maximizes σ_ℓ^2 produces the desired result, after possibly passing to the subsequence $\{j_{2k}\}$ or $\{j_{2k+1}\}$. $\square$

5. Exceptional GAFs

In this section, we construct GAFs with exceptional properties. In particular, we show the strict inclusions in (7).

5A. *H^2–Bloch GAFs are not always BMO GAFs.* Both lacunary and regularly varying H^2-GAFs are VMOA. Sledd [1981, Theorem 3.5] constructs an example of an H^2 random series that is not Bloch, and so is not BMOA. This leaves open the possibility that once an H^2-GAF is Bloch, it additionally is BMO. We give an example that shows there are H^2-GAFs that are Bloch but not BMO.

Recall (3), that for an interval $I \subseteq \mathbb{R}/\mathbb{Z}$, any $p \geq 1$, and any $\mathrm{L}^p(\mathbb{T})$ function f,

$$M_I^p(f) := \fint_I \left| f(e(\theta)) - \fint_I f \right|^p \mathrm{d}\theta, \quad \text{where } \fint_I f(e(\theta))\,\mathrm{d}\theta := \frac{1}{|I|}\int_I f(e(\theta))\,\mathrm{d}\theta.$$

Lemma 5.1. *For every $R > 0$, there exists $n_0 = n_0(R)$ such that for any $n > n_0$ there is a polynomial $f(z) := \sum_{k=n}^m a_k \xi_k z^k$ with the following properties*:

(i) $\sum_k a_k^2 = 1$.

(ii) $\mathbb{E}\|f\|_* \geq R$.

(iii) $\mathbb{E}\|f\|_{\mathcal{B}} \leq C$, *where $C > 0$ is an absolute constant.*

We can then use this lemma to construct the desired GAF.

Theorem 5.2. *There exists an H^2, Bloch, non-*BMOA *GAF.*

Proof. Let $\{\beta_i\}$ and $\{R_i\}$ be two positive sequences with $\{\beta_i\} \in \ell_1$ and $\beta_i R_i \to \infty$. Let f_i be a sequence of independent random Gaussian polynomials given by Lemma 5.1 having

$$\mathbb{E}\|f_i\|_* \geq R_i \quad \text{and} \quad \mathbb{E}\|f_i\|_{\mathcal{B}} \leq C.$$

The function $f = \sum_i \beta_i f_i$ satisfies, for all $\theta \in \mathbb{R}/\mathbb{Z}$,

$$\mathbb{E}|f(e(\theta))|^2 = \sum_{i=1}^{\infty} \beta_i^2 < \infty,$$

and so f is in L^2. The Bloch norm satisfies

$$\mathbb{E}\|f\|_{\mathcal{B}} \leq \sum_{i=1}^{\infty} \mathbb{E}\beta_i \|f_i\|_{\mathcal{B}} < \infty.$$

Finally, by the contraction principle (Proposition 2.1),

$$\mathbb{E}\|f\|_* \geq \beta_i \mathbb{E}\|f_i\|_* \geq \beta_i R_i \to \infty,$$

as $i \to \infty$. Therefore $\|f\|_* = \infty$ a.s. by Corollary 2.8. □

Remark 5.3. It is possible to choose the polynomial $\{f_i\}$ to have disjoint Fourier support, although it is not necessary for the proof, as we have picked them to be independent.

***Proof of Lemma 5.1*.**

Construction of f. Let $r \in \mathbb{N}$ be some parameter to be fixed later (sufficiently large). Let

$$\{\lambda_{i,j} : i, j \in \{1, 2, \dots, r\}\} \cup \{1\}$$

be real numbers that are linearly independent over the rationals and that satisfy

$$\lambda_{i,j} \in [2^i, 2^i + 4^{-r}]. \tag{34}$$

By Kronecker's theorem, for every $\omega \in \left\{0, \frac{1}{2}\right\}^{r\times r}$ there is an $m = m(\omega)$ such that

$$|\{m\lambda_{i,j}\} - \omega_{i,j}| \le 4^{-r} \quad \text{for all } i, j = 1, \dots, r, \tag{35}$$

where as usual $\{x\} = x - \lfloor x \rfloor$ is the fractional value.

Let $n_0 = 4^r(\max_\omega m(\omega) + 1)$, and let $n > n_0$ be arbitrary. Define

$$a_k = \begin{cases} \frac{1}{r} & \text{if } k = \lfloor n\lambda_{i,j} \rfloor \text{ for some } i, j = 1, \dots, r, \\ 0 & \text{otherwise.} \end{cases}$$

For brevity, write $\zeta_{i,j} = \xi_{\lfloor n\lambda_{i,j} \rfloor}$ for any $i, j = 1, \dots, r$. Note that the $\zeta_{i,j}$ are independent and we can write

$$f(z) = \frac{1}{r} \sum_{i,j=1}^{r} \zeta_{i,j} z^{\lfloor n\lambda_{i,j} \rfloor}. \tag{36}$$

Lower bound for $\mathbb{E}\|f\|_*$. Define a random variable $\omega \in \left\{0, \frac{1}{2}\right\}^{r\times r}$ by

$$\omega_{i,j} = \begin{cases} 0 & \text{if } \mathrm{Re}\zeta_{i,j} \ge 0, \\ \frac{1}{2} & \text{if } \mathrm{Re}\zeta_{i,j} < 0. \end{cases}$$

Let I be the interval of length $1/n$ centered around $m(\omega)/n$.

We will show that $\mathbb{E} M_I^2(f)$ is large. To do so, we give an effective approximation for $\mathrm{Re} f$ on I. Define

$$g(\theta) := \sum_{i=1}^{r} \Xi_i \cos(2\pi \cdot 2^i n\theta) \quad \text{where } \Xi_i := \frac{1}{r} \sum_{j=1}^{r} |\mathrm{Re}\zeta_{i,j}|.$$

Notice that g is $1/n$-periodic and therefore

$$M_I^2(g) = \fint_I |g(\theta)|^2 \, d\theta = \frac{1}{2} \sum_{i=1}^{r} \Xi_i^2.$$

Hence $\mathbb{E} m_I^2(g) \ge Cr$ for some absolute constant $C > 0$, and so it remains to approximate f by g.

Claim 5.4. *There is a sine trigonometric polynomial h such that with*

$$E = E(\theta) := \mathrm{Re} f\left(e\left(\frac{m(\omega)}{n} + \theta\right)\right) - g(\theta) - h(n\theta)$$

and, for $|\theta| \le 1/n$,

$$|E(\theta)| \le 3 \cdot 4^{-r} \sum_{i,j} |\zeta_{i,j}|.$$

Proof. By (35),

$$d\Big(\frac{m(\omega)}{n}\lfloor n\lambda_{i,j}\rfloor-\omega_{i,j},\mathbb{Z}\Big)\le 4^{-r}+\frac{m(\omega)}{n}\le 2\cdot 4^{-r}.$$

By (34), $\lfloor n\lambda_{i,j}\rfloor\in[n2^i,n2^i+n4^{-r}]$, and so for $|\theta|\le 1/n$,

$$|\theta\lfloor n\lambda_{i,j}\rfloor-\theta n2^i|\le 4^{-r}.$$

Combining these two estimates, for $|\theta|\le 1/n$ and for all $i,\, j=1,2,\ldots,r$,

$$\left|\operatorname{Re}\Big(\zeta_{i,j}e\Big(\Big(\frac{m(\omega)}{n}+\theta\Big)\lfloor n\lambda_{i,j}\rfloor\Big)\Big)-\operatorname{Re}(\zeta_{i,j}e(\omega_{i,j}+2^in\theta))\right|\le 3\cdot 4^{-r}|\zeta_{i,j}|.$$

Using that $e\big(\theta+\frac{1}{2}\big)=-e(\theta)$, the claim follows by applying the previous estimate term by term to (36). □

We now bound the oscillation $M_I^2(f)$ as follows. Using $⨍_I g=⨍_I h=0$ and the orthogonality of g and h on I,

$$\begin{aligned}\left[⨍_I\left|f(e(\theta))-⨍_I f\right|^2d\theta\right]^{1/2}&\ge\left[⨍_I\left|\operatorname{Re}f(e(\theta))-⨍_I\operatorname{Re}f\right|^2d\theta\right]^{1/2}\\&\ge\left[⨍_I|g(\theta)+h(\theta)|^2\,d\theta\right]^{1/2}-\left[⨍_I|E(\theta)|^2\,d\theta\right]^{1/2}\\&\ge\left[⨍_I|g(\theta)|^2\,d\theta\right]^{1/2}-3\cdot 4^{-r}\sum_{i,j}|\zeta_{i,j}|\\&\ge\left[\frac{1}{2}\sum_{i=1}^r\Xi_i^2\right]^{1/2}-3\cdot 4^{-r}\sum_{i,j}|\zeta_{i,j}|\end{aligned}$$

Using [Girela 2001, Proposition 4.1], there is a constant $C_2\ge 1$ such that

$$\mathbb{E}\|f\|_*\ge\frac{1}{C_2}\mathbb{E}\left(\left[⨍_I\left|f(e(\theta))-⨍_I f\right|^2d\theta\right]^{1/2}\right)\ge\frac{\sqrt{r}}{C}-C\cdot 4^{-r}r^2$$

for some sufficiently large absolute constant $C>0$.

Upper bound for $\mathbb{E}\|f\|_{\mathcal{B}}$. We begin by computing

$$f'(z)=\frac{1}{r}\sum_{i,j}\zeta_{i,j}\lfloor n\lambda_{i,j}\rfloor z^{\lfloor n\lambda_{i,j}\rfloor-1},$$

with the sum over all $i,\, j$ in $\{1,2,\ldots,r\}$. Let $\Theta_i=\frac{1}{r}\sum_j|\zeta_{i,j}|$. Then, for $t=|z|<1$,

$$\begin{aligned}(1-|z|)|f'(z)|\le\frac{1-t}{r}\sum_{i,j}|\zeta_{i,j}|n2^{i+1}t^{n2^i-1}&\le\big(\max_i\Theta_i\big)(1-t)\sum_i n2^{i+1}t^{n2^i-1}\\&\le 2\big(\max_i\Theta_i\big)(1-t)\sum_{i=1}^\infty t^{i-1}\le 2\big(\max_i\Theta_i\big).\end{aligned}$$

Using the Pythagorean property of subgaussian norms [Vershynin 2018, Proposition 2.6.1], the random variables Θ_i have subgaussian norm $C/\sqrt{i}$, and hence using standard estimates,

$$\sup_{r\in\mathbb{N}}\Big[\mathbb{E}\max_{i=1,2,\dots,r}\Theta_i\Big]<\infty.$$

This concluded the proof of Lemma 5.1. $\square$

5B. ***BMO GAFs are not always VMO GAFs.*** We answer a question of [Sledd 1981], showing that there are GAFs which are in BMOA but not in VMOA. We begin by defining a new seminorm on BMOA

$$\|f\|_{*,n}:=\sup_{I:2^{-n}\le|I|\le 2^{-(n-1)}} M_I^1(f),$$

where the supremum is over intervals $I\subset\mathbb{R}/\mathbb{Z}$.

Lemma 5.5. *There is a constant $c>0$ and an $m\in\mathbb{N}$ such that, for all integers $n\ge m$ and for all polynomials p with coefficients supported in* $[2^n,2^{n+1}]$,

$$\|p\|_*\ge\|p\|_{*,n-m}\ge c\|p\|_\infty\ge c\|p\|_*.$$

Proof. The first inequality is trivial. The last inequality is [Girela 2001, Proposition 2.1]. Thus it only remains to prove the second inequality. Recall T_n, the dyadic trapezoidal kernel from (15), which satisfies for all $n\in\mathbb{N}$ that

$$\hat{T}_n(j)=1\quad\text{for } j\in[2^n,2^{n+1}],\qquad \hat{T}_n(0)=0,\qquad\text{and}\qquad \|T_n\|_\infty\le 10\cdot 2^n$$

(see (14) and (15) — this follows by bounding $\|K_n\|_\infty=n+1$ and using the positivity of K). From the condition on the support of the coefficients, $p\star T_n=p$. As the constant coefficient of T_n vanishes, $1\star T_n=0$, and therefore we have the identity that for any $I\subseteq\mathbb{R}/\mathbb{Z}$ and any $\phi\in\mathbb{R}/\mathbb{Z}$,

$$\begin{aligned}p(e(\phi))&=\bigg(\Big(p-⨍_I p\Big)\star T_n\bigg)(e(\phi))\\&=\int_0^1\Big(p(e(\theta))-⨍_I p\Big)T_n(e(\phi-\theta))\,d\theta.\end{aligned}\tag{37}$$

Fix $m\in\mathbb{N}$. Let I be the interval around ϕ of length $2\cdot 2^{m-n}$. Then, for $n\ge m+1$,

$$\begin{aligned}|p(e(\phi))|&\le\int_{I\cup I^c}\Big|p(e(\theta))-⨍_I p\Big|\,|T_n(e(\phi-\theta))|\,d\theta\\&\le\|T_n\|_\infty\int_I\Big|p(e(\theta))-⨍_I p\Big|\,d\theta+2\|p\|_\infty\int_{I^c}|T_n(e(\phi-\theta))|\,d\theta.\end{aligned}\tag{38}$$

The first summand we control as follows (using the length of $|I|$ and $\|T_n\|_\infty\le 10\cdot 2^n$):

$$\begin{aligned}\|T_n\|_\infty\int_I\Big|p(e(\theta))-⨍_I p\Big|\,d\theta&\le 20\cdot 2^m⨍_I\Big|p(e(\theta))-⨍_I p\Big|\,d\theta\\&\le 20\cdot 2^m\|p\|_{*,n-m-1}.\end{aligned}\tag{39}$$

For the second summand, using (16),

$$
\begin{aligned}
2\|p\|_\infty \int_{I^c} |T_n(e(\phi-\theta))|\,d\theta &\le 4\|p\|_\infty \int_{2^{m-n}}^{1/2} |T_n(e(\theta))|\,d\theta \\
&\le 80\cdot 2^{-n}\|p\|_\infty \int_{2^{m-n}}^{1/2} |1-e(\theta)|^{-2}\,d\theta \\
&\le 20\cdot 2^{-n}\|p\|_\infty \int_{2^{m-n}}^{\infty} \theta^{-2}\,d\theta \\
&\le 20\cdot 2^{-m}\|p\|_\infty. \qquad (40)
\end{aligned}
$$

Applying (39) and (40) to (38),

$$
\|p\|_\infty \le 20\cdot 2^{m}\|p\|_{*,n-m-1} + 20\cdot 2^{-m}\|p\|_\infty.
$$

Taking $m=5$, we conclude

$$
\|p\|_\infty \le 2^{11}\|p\|_{*,n-6}. \qquad \square
$$

The previous lemma allows us to estimate $\|\cdot\|_*$ for polynomials supported on dyadic blocks efficiently in terms of the supremum norm. Hence, we record the following simple observation.

Lemma 5.6. *For any $n\ge 2$, let f_n be the Gaussian polynomial*

$$
f_n(z) = \frac{1}{\sqrt{n\log n}} \sum_{k=n}^{2n-1} \xi_k z^k.
$$

Then there is an absolute constant $C>0$ such that

$$
C^{-1} < \mathbb{E}\|f_n\|_\infty < C.
$$

Further, for all $t\ge 0$,

$$
\mathbb{P}[|\|f_n\|_\infty - \mathbb{E}\|f_n\|_\infty| > t] \le 2e^{-(\log n)t^2}.
$$

Proof. Observe that the family $\{f_n(e(k/n)) : 0\le k<n\}$ consists of i.i.d. complex Gaussian random variables of variance $1/\log n$. Hence,

$$
\mathbb{E}\|f_n\|_\infty \ge \mathbb{E}\max_{0\le k<n} |f_n(e(k/n))| \ge C
$$

for some constant $C>0$ (see [Vershynin 2018, Exercise 2.5.11]). Conversely, there is an absolute constant such that for any polynomial p of degree $2n$ (e.g., see [Rakhmanov and Shekhtman 2006]),

$$
\|p\|_\infty \le C \max_{0\le k\le 4n} |p(e(k/(4n)))|.
$$

Hence using that each $f_n(e(k/(4n)))$ is complex Gaussian of variance $1/\log n$, we conclude that there is another constant $C>0$ so that

$$
\mathbb{E}\|f_n\|_\infty \le C
$$

(see [Vershynin 2018, Exercise 2.5.10]). The concentration is a direct consequence of Proposition 2.2. $\square$

Let $\{n_k\}$ be a monotonically increasing sequence of positive integers, to be chosen later. Let f_k be independent Gaussian polynomials as in Lemma 5.6 with $n = 2^{n_k}$. Let $\{a_k\}$ be a nonnegative sequence. Define $g = \sum_{k=1}^{\infty} a_k f_k$. Under the condition that $\sum_{k=1}^{\infty} a_k^2/n_k < \infty$, we have that g is an H^2-GAF.

Lemma 5.7. *Let $n_1 = 1$ and define $n_{k+1} = 3^{n_k}$ for all $k \geq 0$. If the sequence $\{a_k\}$ is bounded, then g is in* BMOA *almost surely. Furthermore, if $\lim_{k\to\infty} a_k = 0$, then g is in* VMOA *almost surely.*

Proof. Without loss of generality we may assume all $a_k \leq 1$. Observe that

$$\|g\|_* = \sup_{\ell\in\mathbb{N}} \|g\|_{*,\ell}.$$

Therefore, if $\sup_{\ell\in\mathbb{N}} \|g\|_{*,\ell} < \infty$ a.s., then g is in BMOA. If, furthermore, $\lim_{\ell\to\infty} \|g\|_{*,\ell} = 0$ a.s., then g is in VMOA almost surely.

Put $g_j = a_j f_j$ for all $j \in \mathbb{N}$. Fix $\ell \in \mathbb{N}$ and let k be such that $n_{k-1} - m \leq \ell \leq n_k - m$, where m is the constant from Lemma 5.5, and take the decomposition $g = g_{<k-1} + g_{k-1} + g_k + g_{>k}$. Then

$$\|g_{>k}\|_{*,\ell} \leq 2^{\ell/2}\|g_{>k}\|_2 \leq 2^{n_k/2}\|g_{>k}\|_2,$$

which follows from Cauchy–Schwarz applied to $M_I^1(g_{>k})$ for an interval $|I| \geq 2^{-\ell}$. On the other hand,

$$\|g_{<k}\|_{*,\ell} \leq 2^{-\ell+1}\|g'_{<k}\|_\infty \leq 2^{-\ell+1}2^{n_{k-2}+1}\|g_{<k}\|_\infty \leq 2^{-n_{k-1}+n_{k-2}+m+2}\|g_{<k}\|_\infty,$$

where the penultimate inequality is Bernstein's inequality for polynomials. We conclude that

$$\|g\|_{*,\ell} \leq 2^{-n_{k-1}+n_{k-2}+m+2}\|g_{<k}\|_\infty + 2\|g_k\|_\infty + 2\|g_{k-1}\|_\infty + 2^{n_k/2}\|g_{>k}\|_2.$$

Using Lemma 5.6 and Borel–Cantelli,

$$D := \sup_k \|f_k\|_\infty < \infty \quad \text{a.s.}$$

In particular,

$$\|g_{<k}\|_\infty \leq kD.$$

Meanwhile, the family $\{\|f_j\|_2^2 \cdot 2 \cdot 2^{n_j}\log(2^{n_j})\}$ consists of independent χ^2 random variables with 2^{n_j+1} degrees of freedom, respectively. Hence,

$$R := \sup_j\{\|g_j\|_2\sqrt{n_j}\} < \infty \quad \text{a.s.}$$

Therefore,

$$\|g_{>k}\|_2^2 = \sum_{j>k}\|g_j\|_2^2 \leq R^2\sum_{j>k}\frac{1}{n_j} \leq \frac{3R^2}{n_{k+1}}.$$

Finally, we have

$$\|g\|_{*,\ell} \leq 2^{-n_{k-1}+n_{k-2}+m+2}kD + 2(a_{k-1}+a_k)D + \sqrt{3}\cdot 2^{n_k/2}\frac{R}{\sqrt{n_{k+1}}}.$$

By our choice of n_k (recalling $k = k(\ell)$), the last expression is *uniformly* bounded in ℓ almost surely. In addition, if $a_k \to 0$, then

$$\limsup_{\ell\to\infty}\|g\|_{*,\ell} \leq \limsup_{k\to\infty} 2(a_{k-1}+a_k)D = 0. \qquad \square$$

Remark 5.8. A more careful analysis of $\|f_{>k}\|_{*,\ell}$ reveals that it suffices to have $n_{k+1}/n_k > c > 1$ for some c to bound $\|f_{>k}\|_{*,\ell}$ uniformly over all ℓ. We will not pursue this here.

We now turn to proving the existence of the desired GAF.

Theorem 5.9. *There is a BMO GAF which is almost surely not a VMO GAF.*

Proof. We let g be as in Lemma 5.7 with $a_k = 1$ for all k. By making t sufficiently small and using the contraction principle (Proposition 2.1), Lemma 5.5 and Lemma 5.6, for all $k \in \mathbb{N}$,

$$2\mathbb{P}(\|g\|_{*,n_k-m} > t) \geq \mathbb{P}(\|f_k\|_{*,n_k-m} > t) \geq \mathbb{P}(\|f_k\|_\infty > ct) \geq \tfrac{1}{2}.$$

Therefore by the reverse Fatou lemma,

$$\mathbb{P}\left(\limsup_{k\to\infty} \|g\|_{*,n_k-m} > t\right) \geq \limsup_{k\to\infty} \mathbb{P}(\|g\|_{*,n_k-m} > t) \geq \tfrac{1}{4}.$$

This implies, by Proposition 2.9, that g is not in VMOA a.s. $\square$

Finally, we show there is a VMO GAF which is not Sledd.

Lemma 5.10. *There is an absolute constant $c > 0$ such that for all $\epsilon > 0$ there is an $n_0(\epsilon)$ sufficiently large such that, for all $n \geq n_0$ and for all intervals $I \subset \mathbb{R}/\mathbb{Z}$ with $|I| = \epsilon$,*

$$\mathbb{P}\left(\textit{there exists } J \subset I \textit{ an interval with } |J| = c/n \textit{ such that } \min_{x\in J} |f_n(x)| > \tfrac{1}{4}\right) \geq \tfrac{1}{3},$$

where f_n is as in Lemma 5.6.

Proof. We again use the observation that the family $\{f_n(e(k/n)) : 0 \leq k < n\}$ consists of i.i.d. complex Gaussian random variables of variance $1/\log n$. Let I be an interval as in the statement of the lemma. Let I' be the middle third of that interval. Then for any n, there are at least $n\epsilon/4$ many k such that k/n are in I'. For any such k and any t,

$$\mathbb{P}[|f_n(e(k/n))| > t] = e^{-(\log n)t^2}.$$

Hence, if we define n_0 so that $n_0^{3/4}\epsilon = 4\log(3)$, then for all $n \geq n_0$,

$$\mathbb{P}\left[\text{for all } k : k/n \in I', |f_n(e(k/n))| \leq \tfrac{1}{2}\right] \leq e^{-n^{3/4}\epsilon/4} \leq \tfrac{1}{3}.$$

Using Bernstein's inequality and Lemma 5.6, there is an absolute constant such that

$$\|f_n'\|_\infty \leq 2n\|f_n\|_\infty \leq Cn,$$

except with probability $1/n$. Hence, if we let J be the interval of length c/n around a point in I' where $|f_n(e(k/n))| > \frac{1}{2}$, then $\min_{x\in J}|f_n(x)| \geq \frac{1}{4}$ except with probability $\frac{2}{3}$. $\square$

Theorem 5.11. *There exists a GAF that is almost surely in* VMOA *and which is almost surely not Sledd.*

Proof. We let g be as in Lemma 5.7 with $a_k \to 0$ to be defined, so that g is almost surely in VMOA.

We define a nested sequence of random intervals $\{J_\ell\}$. Let $J_0 = \mathbb{R}/\mathbb{Z}$. Define a subsequence n_{k_ℓ} inductively by letting n_{k_ℓ} be the smallest integer bigger than $n_0(c/n_{k_{\ell-1}})$ for $\ell > 1$ and $n_0(c)$ for $\ell = 1$. Let $a_{n_{k_\ell}} = 1/\sqrt{\ell}$, and let $a_j = 0$ if j is not in $\{n_{k_\ell}\}$.

We say that an interval J_ℓ *succeeds* if there is a subinterval J' of length $c/n_{k_{\ell-1}}$ such that $\min_{x\in J'}|f_{k_\ell}| > \frac{1}{4}$. If the interval J_ℓ succeeds, we let $J_{\ell+1} = J'$, and otherwise we let $J_{\ell+1}$ be the interval of length $c/n_{k_{\ell-1}}$ that shares a left endpoint with J_ℓ. The nested intervals $\bar{J}_\ell$ decrease to a point x, and

$$\|f\|_{S(R)}^2 \geq \sum_{\ell=1}^{\infty} \frac{1}{\ell}|f_{k_\ell}(x)|^2 \geq \sum_{\ell=1}^{\infty} \frac{1}{16\ell}\mathbf{1}\{J_\ell \text{ succeeds}\}.$$

From Lemma 5.10, the family $\{\mathbf{1}\{J_\ell \text{ succeeds}\}\}$ consists of independent Bernoulli random variables with parameter at least $\frac{1}{3}$. Then by [Kahane 1985, Chapter 3, Theorem 6], this series is almost surely infinite. □

Acknowledgements

We are very thankful to Gady Kozma, to whom the proof of Lemma 5.1 is due. We are also very thankful to the referees who provided lots of useful feedback that improved the clarity and correctness of the manuscript. Most of the work was performed during mutual visits of the authors, which were supported by a grant from the United States–Israel Binational Science Foundation (BSF Start-Up grant 2018341). Nishry was supported in part by ISF grant 1903/18. Support also provided by European Research Council Advanced Grant 692616 and Starting Grant 678520.

References

[Adamczak 2010] R. Adamczak, "A few remarks on the operator norm of random Toeplitz matrices", *J. Theoret. Probab.* **23**:1 (2010), 85–108. MR Zbl

[Anderson et al. 1974] J. M. Anderson, J. Clunie, and C. Pommerenke, "On Bloch functions and normal functions", *J. Reine Angew. Math.* **270** (1974), 12–37. MR Zbl

[Billard 1963] P. Billard, "Séries de Fourier aléatoirement bornées, continues, uniformément convergentes", *Studia Math.* **22** (1963), 309–329. MR Zbl

[Bryc et al. 2006] W. Bryc, A. Dembo, and T. Jiang, "Spectral measure of large random Hankel, Markov and Toeplitz matrices", *Ann. Probab.* **34**:1 (2006), 1–38. MR Zbl

[Duren 1970] P. L. Duren, *Theory of H^p spaces*, Pure Appl. Math. **38**, Academic, New York, 1970. MR Zbl

[Fefferman and Stein 1972] C. Fefferman and E. M. Stein, "H^p spaces of several variables", *Acta Math.* **129**:3-4 (1972), 137–193. MR Zbl

[Gao 2000] F. Gao, "A characterization of random Bloch functions", *J. Math. Anal. Appl.* **252**:2 (2000), 959–966. MR Zbl

[Girela 2001] D. Girela, "Analytic functions of bounded mean oscillation", pp. 61–170 in *Complex function spaces* (Mekrijärvi, Finland, 1999), edited by R. Aulaskari, Univ. Joensuu Dept. Math. Rep. Ser. **4**, Univ. Joensuu, Finland, 2001. MR Zbl

[Holland and Walsh 1986] F. Holland and D. Walsh, "Criteria for membership of Bloch space and its subspace, BMOA", *Math. Ann.* **273**:2 (1986), 317–335. MR Zbl

[Hough et al. 2009] J. B. Hough, M. Krishnapur, Y. Peres, and B. Virág, *Zeros of Gaussian analytic functions and determinantal point processes*, Univ. Lect. Ser. **51**, Amer. Math. Soc., Providence, RI, 2009. MR Zbl

[Kahane 1985] J.-P. Kahane, *Some random series of functions*, 2nd ed., Cambridge Stud. Adv. Math. **5**, Cambridge Univ. Press, 1985. MR Zbl

[Ledoux 1996] M. Ledoux, "Isoperimetry and Gaussian analysis", pp. 165–294 in *Lectures on probability theory and statistics* (Saint-Flour, France, 1994), edited by P. Bernard, Lecture Notes in Math. **1648**, Springer, 1996. MR Zbl

[Ledoux and Talagrand 1991] M. Ledoux and M. Talagrand, *Probability in Banach spaces: isoperimetry and processes*, Ergebnisse der Mathematik (3) **23**, Springer, 1991. MR Zbl [de Leeuw et al. 1977] K. de Leeuw, Y. Katznelson, and J.-P. Kahane, "Sur les coefficients de Fourier des fonctions continues", *C. R. Acad. Sci. Paris Sér. A-B* **285**:16 (1977), A1001–A1003. MR Zbl

[Marcus and Pisier 1978] M. B. Marcus and G. Pisier, *Necessary and sufficient conditions for the uniform convergence of random trigonometric series*, Lect. Not. Ser. **50**, Aarhus Univ., Denmark, 1978. MR Zbl

[Marczewski and Sikorski 1948] E. Marczewski and R. Sikorski, "Measures in non-separable metric spaces", *Colloq. Math.* **1** (1948), 133–139. MR Zbl

[Meckes 2007] M. W. Meckes, "On the spectral norm of a random Toeplitz matrix", *Electron. Comm. Probab.* **12** (2007), 315–325. MR Zbl

[Nazarov and Sodin 2010] F. Nazarov and M. Sodin, "Random complex zeroes and random nodal lines", pp. 1450–1484 in *Proceedings of the International Congress of Mathematicians*, *III* (Hyderabad, 2010), edited by R. Bhatia et al., Hindustan, New Delhi, 2010. MR Zbl

[Nekrutkin 2013] V. V. Nekrutkin, "Remark on the norm of random Hankel matrices", *Vestnik St. Petersburg Univ. Math.* **46**:4 (2013), 189–192. MR Zbl

[Paley and Zygmund 1930] R. E. A. C. Paley and A. Zygmund, "On some series of functions, I", *Math. Proc. Cambridge Philos. Soc.* **26**:3 (1930), 337–357. Zbl

[Peller 2003] V. V. Peller, *Hankel operators and their applications*, Springer, 2003. MR Zbl

[Rakhmanov and Shekhtman 2006] E. Rakhmanov and B. Shekhtman, "On discrete norms of polynomials", *J. Approx. Theory* **139**:1-2 (2006), 2–7. MR Zbl

[Sen and Virág 2013] A. Sen and B. Virág, "The top eigenvalue of the random Toeplitz matrix and the sine kernel", *Ann. Probab.* **41**:6 (2013), 4050–4079. MR Zbl

[Sledd 1981] W. T. Sledd, "Random series which are BMO or Bloch", *Michigan Math. J.* **28**:3 (1981), 259–266. MR Zbl

[Stein 1966] E. Stein, "Classes H^p, multiplicateurs et fonctions de Littlewood–Paley: applications de résultats antérieurs", *C. R. Acad. Sci. Paris Sér. A-B* **263** (1966), A780–A781. MR Zbl

[Talagrand 2001] M. Talagrand, "Majorizing measures without measures", *Ann. Probab.* **29**:1 (2001), 411–417. MR Zbl

[Vershynin 2018] R. Vershynin, *High-dimensional probability: an introduction with applications in data science*, Cambridge Ser. Stat. Probab. Math. **47**, Cambridge Univ. Press, 2018. MR Zbl

[Wulan 1994] H. Wulan, "Random power series with bounded mean oscillation characteristics", *Acta Math. Sci.* (*Chinese*) **14**:4 (1994), 451–457. MR Zbl

Received 8 Feb 2020. Revised 25 Feb 2021. Accepted 4 May 2021.

ALON NISHRY: alonish@tauex.tau.ac.il
School of Mathematical Sciences, Department of Pure Mathematics, Tel Aviv University, Tel Aviv, Israel

ELLIOT PAQUETTE: elliot.paquette@gmail.com
Department of Mathematics, McGill University, Montreal, Canada

ANALYSIS AND PDE
Vol. 16 (2023), No. 1, pp. 119–171
DOI: 10.2140/apde.2023.16.119

msp

GENERIC KAM HAMILTONIANS ARE NOT QUANTUM ERGODIC

SEÁN GOMES

We show that under generic conditions, the quantisation of a 1-parameter family of KAM perturbations $P(x, \xi; t)$ of a completely integrable and Kolmogorov nondegenerate Gevrey smooth Hamiltonian is not quantum ergodic for a full-measure subset of parameter values $t \in (0, \delta)$.

1. Introduction

1A. *Hamiltonian dynamics.* Let M be a compact boundaryless Riemannian G^ρ smooth manifold of dimension $n \geq 2$, and let $P(x, \xi) \in \mathcal{C}^\infty(T^*M)$ be a completely integrable Hamiltonian with $P(x, \xi) \to \infty$ as $|\xi| \to \infty$. Complete integrability is the assumption that there exist n functionally independent conserved quantities of the Hamiltonian flow that are pairwise in involution.

The Liouville–Arnold theorem asserts that we can locally choose symplectomorphisms

$$\chi : \mathbb{T}^n \times D \to T^*M \tag{1A.1}$$

such that the transformed Hamiltonian

$$H^0(\theta, I) = (P \circ \chi)(\theta, I) \tag{1A.2}$$

is independent of θ. It follows that the Hamiltonian flow is quasiperiodic and constrained to n-dimensional Lagrangian tori, given in local coordinates by

$$\dot{I} = 0, \quad \dot{\theta} = \nabla H(I). \tag{1A.3}$$

Under the Kolmogorov nondegeneracy condition $\det(\nabla_I^2 H) \neq 0$, we can locally index the invariant Lagrangian tori Λ_ω by the frequency $\omega = \nabla_I H$ of their quasiperiodic motion.

Gomes was supported by an Australian postgraduate award and research training program stipend.
MSC2010: 58J51.
Keywords: quantum ergodicity, KAM Hamiltonians.

If we now consider a smooth one-parameter family of perturbed Hamiltonians given by $H(\theta, I; t)$ in action-angle coordinates with $H(\theta, I; 0) = H^0(I)$, a natural question is whether or not an invariant tori Λ_ω with quasiperiodic motion of frequency ω still exists in the perturbed Hamiltonian dynamics.

This question was resolved positively by Kolmogorov [1954], Arnold [1983], and Moser [1966]. In particular, they established that the Lagrangian invariant tori corresponding to all but an $o(1)$-symplectic-measure subset of frequencies survive this perturbation as the size of the perturbation tends to zero.

In particular persisting tori are those with frequencies ω in a set Ω_κ determined by the Diophantine condition (3B.2), where $\tau > n-1$ is fixed and the choice of κ then dictates the measure of the union of preserved tori.

Popov [2004b] used a local version of the KAM theorem to construct a Birkhoff normal form for Gevrey class Hamiltonians H about Λ. This normal form generalises the notion of "action-angle" variables of a completely integrable Hamiltonian as discussed in [Arnold 1989]. As a consequence of the normal form construction, Popov obtained an effective stability result for the Hamiltonian flow near the union of remaining invariant tori. The natural setting for the estimates is that of Gevrey regularity. This work generalises earlier work in [Popov 2000a; 2000b], where a Birkhoff normal form is constructed for real-analytic Hamiltonians.

1B. *Quantum ergodicity.* We now consider the quantisation of a KAM Hamiltonian system given by a family of self-adjoint and uniformly elliptic semiclassical pseudodifferential operators

$$\mathcal{P}_h(t) = \sum_{j=0}^{m} P_j(x, hD; t)h^j, \tag{1B.1}$$

with real-valued full symbol in the Gevrey class $S_\ell(T^*M)$ from Definition B.5, analytic in the parameter t, where $\ell = (\rho, \mu, \nu)$, with $\rho(\tau+n)+1 > \mu > \rho' = \rho(\tau+1)+1$ and $\nu = \rho(\tau+n+1)$. Furthermore, we assume $\mathcal{P}_h(t)$ acts on half-densities in $\mathcal{C}^\infty(M; \Omega^{1/2})$ with principal symbol $P_0(x, \xi; t)$ completely integrable and nondegenerate at $t = 0$, and with vanishing subprincipal symbol. The operator $\mathcal{P}_h(t)$ then has an orthonormal basis of eigenfunctions $u_j(t; h)$ and corresponding real eigenvalues $E_j(t; h) \to \infty$ for each fixed t, h.

The Bohr correspondence principle asserts that aspects of the classical dynamics should be reflected in the spectral theory of $\mathcal{P}_h(t)$ in the semiclassical limit $h \to 0$. A rigorous manifestation of this correspondence principle is the celebrated quantum ergodicity theorem, due to [Shnirelman 1974; Colin de Verdière 1985; Zelditch 1987], which asserts that billiards with ergodic geodesic flow have eigenfunctions satisfying a quantum notion of equidistribution, made precise using the machinery of pseudodifferential operators.

We work with a semiclassical formulation of quantum ergodicity. Let $d\mu_E$ denote the measure on the energy surface $\Sigma_E = p^{-1}(E)$ induced by the symplectic measure $|d\xi \wedge dx|$ on T^*M by

$$|d\mu_E \wedge dE| = |d\xi \wedge dx|. \tag{1B.2}$$

If a Hamiltonian $p(x, \xi) \in \mathcal{C}^\infty(T^*M)$ generates an ergodic Hamiltonian flow on every energy surface Σ_E with $E \in [a, b]$ and $dp|_{p^{-1}([a,b])} \neq 0$, then for any semiclassical pseudodifferential operator A of semiclassical order 0, the quantum ergodicity theorem states that

$$h^n \sum_{E_j(h)\in[a,b]} \left| \langle A_h u_j(h), u_j(h)\rangle - \frac{1}{\mu_{E_j}(\Sigma_{E_j})} \int_{\Sigma_{E_j}} \sigma(A)\, d\mu_{E_j} \right|^2 \to 0. \tag{1B.3}$$

The quantum ergodicity theorem is originally due to Shnirelman [1974], Zelditch [1987], and Colin de Verdière [1985]. The semiclassical formulation of the quantum ergodicity theorem (1B.3) is a straightforward consequence of the sharper formulation in [Helffer et al. 1987], or [Dyatlov and Guillarmou 2014], in which the statement is localised to $O(h)$ energy bands. From (1B.3), a standard diagonal argument introduced in [Colin de Verdière 1977] shows that

$$\lim_{h\to 0} \left| \langle A_h u_j(h), u_j(h)\rangle - \frac{1}{\mu_{E_j}(\Sigma_{E_j})} \int_{\Sigma_{E_j}} \sigma(A)\, d\mu_{E_j} \right| = 0 \tag{1B.4}$$

uniformly for a family $\Lambda(h) \subset \{E_j(h) \in [a,b]\}$ of full density, in the sense that

$$\frac{\#\Lambda(h)}{\#\{E_j(h)\in[a,b]\}} \to 1. \tag{1B.5}$$

We say that a semiclassical pseudodifferential operator of the form (1B.1) is *quantum ergodic* if its eigenfunctions satisfy (1B.3).

In the appendix to [Marklof and O'Keefe 2005], Zelditch raises the question of *converse quantum ergodicity*: to what extent is it possible for nonergodic Hamiltonian systems such as those in the KAM regime to have quantum ergodic quantisations? In the extreme situation of quantum complete integrability, rigorous results on eigenfunction microlocalisation onto unions of Lagrangian tori have been established in [Toth and Zelditch 2003], which clearly rules out quantum ergodicity. In the intermediate regimes between complete integrability and ergodicity, fewer rigorous results on the question of converse quantum ergodicity are known. In the appendix to [Marklof and O'Keefe 2005], Zelditch shows that the "pimpled spheres", which are S^2 with a metric deformed polar cap, are not quantum ergodic, exploiting the periodicity of the flow in a strong way. In [Gutkin 2009] it is shown that the "racetrack billiard" is quantum ergodic but not ergodic, with phase space splitting into two disjoint invariant sets of equal measure.

As KAM dynamics are far from ergodic dynamics in character, the Bohr correspondence principle suggests that $\mathcal{P}_h(t)$ is typically *not* quantum ergodic, and that under generic conditions on the perturbation, there could exist sequences of eigenfunctions for $\mathcal{P}_h(t)$ with semiclassical mass entirely supported on individual invariant tori.

This localisation was proven for quasimodes in [Popov 2004a], where semiclassical Fourier integral operators were used to construct a quantum Birkhoff normal form for a class of semiclassical pseudodifferential operators $\mathcal{P}_h(t)$. This quantum Birkhoff normal form is used to obtain a family of quasimodes microlocalised near the union of KAM Lagrangian tori of a Hamiltonian associated to $\mathcal{P}_h$. A similar construction was previously made in [Colin de Verdière 1977], which establishes the existence of quasimodes microlocalised near the Lagrangian tori of a completely integrable Hamiltonian on a compact smooth manifold.

As pointed out in [Zelditch 2004], however, the passage from quasimode microlocalisation statements to microlocalisation statements for genuine eigenfunctions typically requires information on the spectral concentration of the operator in question.

One way in which this information can be obtained is by considering the spectral flow of $\mathcal{P}_h(t)$ in an analytic parameter t as in this paper. The Hadamard variational formula allows us to rule out spectral concentration for full measure t, given suitable information on the expectation of the quantum observable

$$\langle \partial_t \mathcal{P}_h(t) u_j(t; h), u_j(t; h)\rangle, \tag{1B.6}$$

which can be obtained from conditions like (1B.4). One can then draw conclusions about eigenfunction microlocalisation from those about quasimode microlocalisation.

In [Hassell 2010], this technique was exploited to obtain the existence of a sequence of Laplacian eigenfunctions on the Bunimovich stadium that does not equidistribute, at least for a full-measure set of aspect ratios. This strategy was also exploited in [Gomes 2018], where the author establishes a weak form of Percival's conjecture for the mushroom billiard.

It is the purpose of this paper to use the same technique to show that quantisations of KAM Hamiltonian systems in the sense of (1B.1) are typically *not* quantum ergodic, at least for full measure $t \in (0, \delta)$.

We follow [Popov 2004a] in working in the category of Gevrey regularity for our Hamiltonian P, due to the availability of explicit and full details of the quantum Birkhoff normal form construction in this setting.

1C. ***Statement of results.*** The following is the main result of this paper.

Theorem 1.1. *Suppose M is a compact boundaryless G^ρ manifold and $\mathcal{P}_h(t)$ is a family of self-adjoint elliptic semiclassical pseudodifferential operators acting on $\mathcal{C}^\infty(M; \Omega^{1/2})$ with fixed positive differential order such that*:

(i) *The operator $\mathcal{P}_h(t)$ has full symbol real-valued, analytic in t, and in the Gevrey class $S_\ell(T^*M)$ from Definition B.5, where $\ell = (\rho, \mu, \nu)$, with $\rho(\tau+n)+1 > \mu > \rho' = \rho(\tau+1)+1$ and $\nu = \rho(\tau+n+1)$.*

(ii) *The principal symbol $P_0(x, \xi; t)$ lies in $G^{\rho,1}(T^*M \times (-1, 1))$.*

(iii) *$P_0(x, \xi; 0)$ is a completely integrable and nondegenerate Hamiltonian.*

(iv) *The subprincipal symbol of $\mathcal{P}_h(t)$ vanishes.*

(v) *In an action-angle variable coordinate patch $\mathbb{T}^n \times D$ for the unperturbed Hamiltonian $P_0(x, \xi; 0)$, the KAM Hamiltonian can be written as $H(\theta, I; t) = P_0(\cdot, \cdot; t) \circ \chi$, and we define $H^0(I) := H(\theta, I; 0)$.*

(vi) *The KAM perturbation is such that*

$$\int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta$$

is nonconstant on some regular energy surface $\{I \in D : H^0(I) = E\}$ in the action-angle coordinate patch.

Then for any regular energy band $P_0^{-1}([a, b])$ with $E \in (a, b)$ for the energy surface in condition (vi), *there exists $\delta > 0$ such that the family of operators $\mathcal{P}_h(t)$ is not quantum ergodic in $[a, b]$ for full measure $t \in (0, \delta)$.*

Remark 1.2. Though we choose to work with Gevrey class Hamiltonians, it should be noted that we only require quasimodes for $\mathcal{P}_h(t)$ of order $O(h^{(3n+2)/2})$ to carry out the arguments in Section 2. In particular this implies that Theorem 1.1 should hold in the $\mathcal{C}^\infty$ setting, where $O(h^\infty)$ quasimodes are constructed in [Colin de Verdière 1977].

Remark 1.3. The condition (vi) is a rather mild one. Indeed for Hamiltonian perturbations of the form $H^0(I)+tH^1(\theta, I)$, it is equivalent to the functional independence of $H^0(I)$ and $\int_{\mathbb{T}^n} H^1(\theta, I)\, d\theta$. This holds for generic choice of H^1.

1D. *Examples.* The broad class of operators satisfying the assumptions of Theorem 1.1 are perturbations of completely integrable Schrödinger-type operators

$$\mathcal{P}_h = -h^2\Delta_g + V(x). \tag{1D.1}$$

In particular, Theorem 1.1 applies to the case of the semiclassical Laplace–Beltrami operator ($V=0$) on a manifold with perturbed metric (M, g_t), where (M, g_0) has completely integrable and nondegenerate geodesic flow.

The model example of a completely integrable geodesic flow is that of the flat torus

$$\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n. \tag{1D.2}$$

The Hamiltonian that generates the geodesic flow on $\mathbb{T}^n$ can be written as $|I|^2$, where $I \in \mathbb{R}^n$ is dual to the spatial variable $\theta \in \mathbb{T}^n$. This is clearly a nondegenerate and completely integrable Hamiltonian system. Similarly, in [Knörrer 1980], it is shown that the geodesic flow on an n-axial ellipsoid E is completely integrable and nondegenerate. Thus the Laplace–Beltrami operator for metric perturbations of both of these manifolds is covered by Theorem 1.1, provided the generic condition (vi) is satisfied.

For an explicit family of examples, one can consider $\mathbb{T}^2$, equipped with the metric

$$g = d\theta_1^2 + d\theta_2^2 + t\chi(\theta_1, \theta_2)\, d\theta_1\, d\theta_2$$

for $t > 0$ small and $\chi \in \mathcal{C}^\infty(\mathbb{T}^2)$ arbitrary. The Hamiltonian corresponding to $-h^2\Delta_g$ is

$$H(\theta, I) = I_1^2 + I_2^2 + t\chi(\theta_1, \theta_2) I_1 I_2,$$

and we have that

$$\int_{\mathbb{T}^2} \partial_t H(\theta, I)\, d\theta = I_1 I_2 \int_{\mathbb{T}^2} \chi(\theta)\, d\theta \tag{1D.3}$$

is nonconstant over any energy surface $|I| = E > 0$; thus condition (vi) of Theorem 1.1 is satisfied.

1E. *Outline of paper.* In Section 3A, we introduce some definitions and notations that are prevalent throughout the paper.

In Section 2, we prove Theorem 1.1 by contradiction. We now outline the strategy of the proof. In Section 2B, under the assumptions of (vi) in Theorem 1.1, Proposition 2.5 makes use of the calculation in Section 3E to obtain an upper bound for the flow speed of a positive density family of the quasieigenvalues constructed in Section 4C. On the other hand, the assumption of quantum ergodicity of $\mathcal{P}_h(t)$ for large measure t yields an estimate for the variation of a large density subset of exact eigenvalues in (2B.22). The results in this section establish a gap (2B.23) between the flow speed of these quasieigenvalues and exact eigenvalues that ensures that individual eigenvalues cannot spend large measure $t \in (0, \delta)$ within $O(h^{n+1})$ distance of any of the quasieigenvalues. This is formalised in Section 2C, where it is deduced that there exists $t_* \in (0, \delta)$ at which there are very few actual eigenvalues within $O(h^{n+1})$ distance of the union of quasieigenvalue windows. An elementary spectral theory contradiction is arrived at from this spectral nonconcentration, completing the proof.

In Section 3, we construct a Gevrey class Birkhoff normal form for the family of Hamiltonians $P(x, \xi; t)$. The construction is that of [Popov 2004b], with our only additional concern being establishing the regularity of this Birkhoff normal form construction in the parameter t. In Section 3E, we compute the derivative of the integrable term $K(I; t)$ of the Birkhoff normal form in the parameter t. This is done by applying two KAM iterations to $P(x, \xi; t)$ prior to the application of the Birkhoff normal form construction of Theorem 3.10.

In Section 4, we recall the quantum Birkhoff normal form construction of [Popov 2004a], formulated in Theorem 4.1. This construction yields a Gevrey family of quasimodes microlocalising on the KAM Lagrangian tori of the Hamiltonian $P(x, \xi; t)$. For the spectral flow arguments in Section 2C we require that the associated quasieigenvalues are smooth in t, which is a statement entirely about the symbols of this quantum Birkhoff normal form.

In Appendix A, we introduce the anisotropic classes of Gevrey functions that are used throughout this paper as well as some of their basic properties.

In Appendix B, we introduce the semiclassical pseudodifferential calculus for Gevrey class symbols.

In Appendix C, we collect two elementary assertions about analytic functions.

In Appendix D, we state and prove a version of the Whitney extension theorem for the anisotropic class of Gevrey functions.

2. Proof of Theorem 1.1

2A. ***Introduction.*** We begin by assuming that $\mathcal{P}_h(t)$ is a family of operators satisfying the assumptions of Theorem 1.1.

Condition (vi) in Theorem 1.1 implies that there exists a nonresonant frequency $\omega_0 \in \widetilde{\Omega}_\kappa$ with associated Lagrangian torus Λ_{ω_0} such that the average of $\partial_t P_0(x, \xi; 0)$ over the torus Λ_{ω_0} differs from the average of $\partial_t P_0(x, \xi; 0)$ over the associated energy surface

$$\{(x, \xi) \in T^*M : P_0(x, \xi; 0) = H^0(I(\omega_0))\}. \tag{2A.1}$$

Moreover, we can ensure that Λ_{ω_0} lies in an arbitrarily small energy window $[a, b]$ about the regular energy E from the condition (vi). Without loss of generality, the hypotheses of Theorem 1.1 thus guarantee the existence of what we shall call a *slow torus*.

Definition 2.1. A *slow torus* in the energy band $[a, b]$ for the unperturbed Hamiltonian

$$H(\theta, I; 0) = H^0(I), \tag{2A.2}$$

written in action-angle coordinates, is a Lagrangian invariant torus Λ_{ω_0} with nonresonant frequency $\omega_0 \in \widetilde{\Omega}_\kappa$ and energy $H^0(I(\omega_0)) \in (a, b)$ in the notation of Theorem 3.10 that satisfies

$$(2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega_0); 0)\, d\theta < \inf_{E \in [a,b]} \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} \partial_t P_0(x, \xi; 0)\, d\mu_E \tag{2A.3}$$

at $t = 0$.

We call such a torus a *slow torus* to draw intuition from the special case where $\partial_t \mathcal{P}_h(t)$ is a positive operator. In this case, as t evolves, the quasieigenvalues associated to such a torus increase as t evolves at a slower rate than the typical increase of eigenvalues at the same energy. The intuition behind this stems from the Hadamard variational formula (2B.8), and the fact that the associated quasimodes microlocalise onto Λ_{ω_0}. This intuition is confirmed in Section 3E, by a more careful analysis of the leading-order behaviour as $t \to 0$ of the integrable term in the Birkhoff normal form established in Theorem 3.10. Under the assumption of quantum ergodicity, this analysis implies a discrepancy (2B.23) in the spectral flow of genuine eigenvalues and quasieigenvalues attached to slow tori. Consequently, we obtain the spectral nonconcentration statement Proposition 2.10.

We begin by using the slow torus condition and choosing $c > 0$ sufficiently small so that

$$(2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega_0; 0); 0)\, d\theta < \inf_{E \in [a,b]} \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} \partial_t P_0(x, \xi; 0)\, d\mu_E - 3c \tag{2A.4}$$

is satisfied.

As the quantum ergodicity condition (1B.3) is preserved upon passing to energy subintervals, we can assume that $[a, b]$ is an arbitrarily small energy window containing $H^0(I(\omega_0; 0))$. In particular, we can scale our interval $[a, b]$ by a small factor λ to ensure that the condition

$$\sup_{E \in [a,b]} \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} \partial_t P_0\, d\mu_E - \inf_{E \in [a,b]} \frac{1}{\mu_E(\Sigma_E)} \int_{\Sigma_E} \partial_t P_0\, d\mu_E =: Q_+(0) - Q_-(0) < \epsilon < c \tag{2A.5}$$

is satisfied for any particular $\epsilon < c$. From the regularity of P_0, one can achieve this by taking

$$\lambda = O(\epsilon). \tag{2A.6}$$

Through the course of this section, we will track the size of various small quantities in terms of this ϵ, which we will eventually take small in the proof of Proposition 2.10.

Proposition 3.14 applies to H, and we obtain a family of symplectomorphisms

$$\chi \in G^{\rho,\rho',\rho'}\big(\mathbb{T}^n \times D \times \big(-\tfrac{1}{2}, \tfrac{1}{2}\big), \mathbb{T}^n \times D\big) \tag{2A.7}$$

and a family of diffeomorphisms

$$\omega \in G^{\rho',\rho'}\big(D \times \big(-\tfrac{1}{2}, \tfrac{1}{2}\big), \Omega\big) \tag{2A.8}$$

such that

$$H(\chi(\theta, I; t); t) = K(I; t) + R(\theta, I; t), \tag{2A.9}$$

where $\partial_I^\alpha R(\theta, I; t) = 0$ for nonresonant actions $I \in E_\kappa(t)$. Using the diffeomorphism (2A.8), we can define an action map $I \in G^{\rho',\rho'}\big(\Omega \times \big(-\tfrac{1}{2}, \tfrac{1}{2}\big)\big)$ implicitly by

$$\tilde{\omega} = \omega(I(\tilde{\omega}; t); t) \tag{2A.10}$$

and we can use this map to specify the action coordinates of a nonresonant torus with fixed frequency at any $t \in \big(-\tfrac{1}{2}, \tfrac{1}{2}\big)$ in the Birkhoff normal form furnished by $\chi(\,\cdot\,, \cdot\,; t)$.

We first obtain a positive-measure family of slow tori near Λ_{ω_0}.

Proposition 2.2. *There exists $r > 0$ and $\delta > 0$ such that for any $\omega \in \overline{\Omega} := B(\omega_0, r) \cap \widetilde{\Omega}_\kappa$, the torus $\Lambda_\omega = \chi(\mathbb{T}^n \times \{I(\omega, t)\})$ has energy*

$$K(I(\omega; t), t) \in [a, b] \tag{2A.11}$$

for all $t \in (0, \delta)$.

In particular, the family of tori

$$\Lambda(t) := \bigcup_{\omega \in \overline{\Omega}} \Lambda_\omega \tag{2A.12}$$

is a positive-measure family of KAM tori entirely contained within the energy band $[a, b]$.

Moreover, r and δ can be chosen small enough to ensure

$$\begin{aligned} (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega; t); t)\, d\theta &< (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega; 0); t)\, d\theta + \epsilon \\ &< \inf_{t \in (0,\delta)} Q_-(t) - 2c \end{aligned} \tag{2A.13}$$

for each $\omega \in \overline{\Omega}$ and each $t \in (0, \delta)$.

We can also choose $\delta > 0$ small enough to ensure that

$$Q_+ - Q_- := \sup_{t \in (0,\delta)} Q_+(t) - \inf_{t \in (0,\delta)} Q_-(t) < 2\epsilon. \tag{2A.14}$$

In particular r, δ can be taken to be $O(\epsilon)$, with constant independent of t and h.

Proof. From the regularity of χ, I, and K established in Theorem 3.10, it follows that we can take $r = O(\lambda)$ to ensure that (2A.11) is satisfied at $t = 0$, where $\lambda = O(\epsilon)$ is as in (2A.6). Similarly, we can ensure that

$$(2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta < (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega_0); t)\, d\theta + \epsilon/2 \tag{2A.15}$$

holds for $|I - I(\omega_0)| = O(\lambda)$. Since (2A.4) is satisfied at $t = 0$, it follows that

$$(2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega; 0); 0)\, d\theta < Q_-(0) - 3c + \epsilon/2 \tag{2A.16}$$

for all $\omega \in \overline{\Omega} = B(\omega_0, r) \cap \widetilde{\Omega}_\kappa$ upon taking $r = O(\lambda)$.

The regularity of χ, I and K in the parameter t then allows us to deduce that (2A.11) and (2A.13) are satisfied for $t \in (0, \delta)$, for sufficiently small $\delta > 0$ and for each $\omega \in \overline{\Omega}$. In particular, we can take $\delta = O(\lambda) = O(\epsilon)$.

Finally, the estimate (2A.14) for small δ follows from the regularity of

$$\frac{1}{\mu_E(\Sigma_E)}\int_{\Sigma_E} \partial_t P_0 \, d\mu_E \tag{2A.17}$$

in t and E. □

We can now apply the quantum Birkhoff normal form construction outlined in Section 4. From Theorem 4.5, we obtain a family of quasimodes that microlocalise onto the family of KAM tori $\Lambda(t)$ introduced in (2A.12).

In particular, following Section 4C, we take $S(t) = \{I(\omega; t) : \omega \in \overline{\Omega}\}$ and define the index set $\mathcal{M}_h(t)$ as in (4C.2). Then for each $m \in \mathcal{M}_h(t) \subset \mathbb{Z}^n$, we have a quasimode v_m with corresponding quasieigenvalue μ_m as in (4C.3). We introduce notation for the union of h^{n+1}-width energy windows about the quasieigenvalue associated to tori in $\Lambda(t)$:

$$W(t; h) := \bigcup_{m \in \mathcal{M}_h(t)} [\mu_m(t; h) - h^{n+1}, \mu_m(t; h) + h^{n+1}]. \tag{2A.18}$$

We also introduce the index set

$$G(h) = \{j \in \mathbb{N} : E_j(t) \in [a, b] \text{ for some } t \in (0, \delta)\} \tag{2A.19}$$

of the eigenvalues that can possibly play a role in the spectral flow considerations in Section 2C.

To conclude this section, we collect asymptotic estimates for the number of eigenvalues and the number of quasieigenvalues that are in the energy window $[a, b]$ as $h \to 0$.

Proposition 2.3. *We have the asymptotic estimate*

$$\#\mathcal{M}_h(t) \sim (2\pi h)^{-n} \mu(\mathbb{T}^n \times \{I(\omega, t) : \omega \in \overline{\Omega}\}) \tag{2A.20}$$

for each $t \in (0, \delta)$.

Furthermore, we have

$$\limsup_{h\to 0} (2\pi h)^n \# G(h) \le \mu(\{(x, \xi) : P_0(x, \xi; 0) \in [a - M\delta, b + M\delta]\}), \tag{2A.21}$$

where M is the uniform bound on spectral flow speed in (2B.11) and $G(h)$ is as in (2A.19).

Here μ denotes the symplectic measure $d\xi\, dx$ on T^*M.

Proof. The estimate (2A.20) is a consequence from (4C.8), and (2A.21) follows from (2B.11) and an application of the semiclassical Weyl law [Zworski 2012, Theorem 14.11]. □

From Proposition 2.3, it follows that we can bound

$$\frac{\# G(h)}{\inf_{t \in (0,\delta)} \#\mathcal{M}_h(t)} \tag{2A.22}$$

for $t \in (0, \delta(\epsilon))$ and $h < h_0(\epsilon)$. Moreover, this upper bound is uniform in ϵ. By the nature of their construction in Proposition 2.2, the quasieigenvalues $\mu_m(t; h)$ lie in $[a, b]$ for all $t \in (0, \delta)$.

It is convenient to introduce the subset $\widetilde{G}(h) \subset G(h)$ given by

$$\widetilde{G}(h) = \{j \in \mathbb{N} : E_j(t) \in [a, b] \text{ for all } t \in (0, \delta)\}. \tag{2A.23}$$

By choosing $\delta(\epsilon) > 0$ appropriately small, we can ensure that a large proportion of eigenvalues that lie in $[a, b]$ for some $t \in (0, \delta)$ lie in $[a, b]$ for all $t \in (0, \delta)$.

Proposition 2.4. *We can choose $\delta(\epsilon) = O(\epsilon^2)$ such that*

$$\frac{\#\widetilde{G}(h)}{\#G(h)} \geq 1 - C\epsilon \tag{2A.24}$$

for all $\epsilon < \epsilon_0$ and $h < h_0(\epsilon)$, where $C > 0$ is a constant.

Proof. We have the bound

$$\frac{\#G(h)}{\#\widetilde{G}(h)} \leq \frac{N_h([a + M\delta, b - M\delta])}{N_h([a - M\delta, b + M\delta])}, \tag{2A.25}$$

where $N_h(I)$ counts the semiclassical eigenvalues of the operator $\mathcal{P}_h(0)$ in I. Recalling that the interval $[a, b]$ is of scale $\lambda = O(\epsilon)$, it follows that for any choice of $\delta = O(\epsilon^2)$, the ratio of phase space volumes

$$\frac{\mu(P_0(x, \xi; 0) \in [a - M\delta, a + M\delta] \cup [b - M\delta, b + M\delta])}{\mu(P_0(x, \xi; 0) \in [a - M\delta, a + M\delta])} \tag{2A.26}$$

can be bounded by a constant multiple of ϵ for all sufficiently small ϵ. Application of the semiclassical Weyl asymptotics to (2A.25) completes the proof. □

2B. ***Eigenvalue and quasieigenvalue variation.*** We now turn our attention to the variation of quasieigenvalues and eigenvalues as $t \in (0, \delta)$ varies. The quasieigenvalues can be handled rather explicitly.

Proposition 2.5. *For any all sufficiently small $\delta(\epsilon) > 0$ and all $t \in (0, \delta)$, we have*

$$\limsup_{h \to 0} \partial_t \mu_m(t; h) \leq Q_- - c \tag{2B.1}$$

for all $m \in \bigcup_{t \in (0,\delta)} \mathcal{M}_h(t)$ uniformly in t.

Proof. From Proposition 3.14, we have

$$K_0(I; t) = H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta + O(t^{9/8}) \tag{2B.2}$$

for any $I \in D$. Hence we have

$$\partial_t (K_0(h(m + \vartheta/4); t)) < (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, h(m + \vartheta/4); 0)\, d\theta + \epsilon \tag{2B.3}$$

for all $t \in (0, \delta(\epsilon))$, taking δ sufficiently small. From the definition of $\mathcal{M}_h(t)$, we know that

$$|h(m + \vartheta/4) - I(\omega; t)| < Lh$$

for some $\omega \in \overline{\Omega}$, and so from the regularity of I in t it follows that

$$\partial_t (K_0(h(m + \vartheta/4); t)) < (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega; t); t)\, d\theta + \epsilon + O(h) \tag{2B.4}$$

for some $\omega \in \overline{\Omega}$. This allows us to compute

$$\begin{aligned}\partial_t \mu_m(t; h) &= \partial_t (K^0(h(m+\vartheta/4); t, h)) \\ &= \partial_t (K_0(h(m+\vartheta/4); t)) + O(h) \\ &< (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I(\omega; t); t)\, d\theta + \epsilon + O(h) \\ &< Q_- - 2c + \epsilon + O(h) \end{aligned} \tag{2B.5}$$

which implies

$$\limsup_{h\to 0} \partial_t \mu_m(t; h) < Q_- - 2c + \epsilon, \tag{2B.6}$$

using (2B.4), (2A.13), and (2A.14). □

In particular, we can choose $B > 0$ and $h_0 > 0$ such that

$$\partial_t \mu_m(t; h) < B < Q_- - c \tag{2B.7}$$

for all $t \in (0, \delta)$ and all $h < h_0$.

Remark 2.6. We abused notation slightly here by writing $\mu_m(t; h)$ even when $m \notin \mathcal{M}_h(t)$. That is, we track the behaviour of $K^0(h(m+\vartheta/4), t; h)$ even for $t \in (0, \delta)$ such that this does not correspond to a quasieigenvalue in our family. This is a necessity due to the rough nature of the set $\{I(\omega; t) : \omega \in \overline{\Omega}\}$ of nonresonant actions. Indices $m \in \mathbb{Z}^n$ will typically be elements of $\mathcal{M}_h(t)$ for only $O(h)$-sized t-intervals at a time.

Remark 2.7. This is the last part of the argument that involves placing an additional restriction on the size of $\delta > 0$.

We now consider the variation of eigenvalues. For each fixed $h > 0$, the operators $\mathcal{P}_h(t)$ comprise an holomorphic family of type A in the sense of [Kato 1966] and so we can choose eigenvalues and corresponding eigenprojections holomorphic in the parameter t. Thus if at each time t we order our eigenpairs $E_j(t; h)$ in order of increasing energy, by holomorphy it follows that E_j will be continuous and piecewise smooth in $(0, \delta)$. On the cofinite set where E_j is differentiable in t, we have

$$\partial_t E_j(t; h) = \langle \partial_t \mathcal{P}_h(t) u_j(t; h), u_j(t; h)\rangle, \tag{2B.8}$$

since (u_j) is an orthonormal basis. We will control (2B.8) using our assumption of quantum ergodicity.

To this end, we now suppose for the sake of contradiction that there exists a positive-measure set $\mathcal{B} \subset (0, \delta)$ such that $\mathcal{P}_h(t)$ is quantum ergodic in the sense of (1B.3) for every $t \in \mathcal{B}$.

Proposition 2.8. *For every $t \in \mathcal{B}$ and $\epsilon > 0$, there exists $h_0(t, \epsilon)$ such that, for all $h < h_0(t, \epsilon)$, we have*

$$\left|\langle \partial_t \mathcal{P}_h(t) u_j(t; h), u_j(t; h)\rangle - \int_{\Sigma_{E_j(t;h)}} \partial_t P_0 \, d\mu_{E_j(t;h)}\right| < \epsilon \tag{2B.9}$$

for a family of indices $S(t; h) \subset \{j \in \mathbb{N} : E_j(t; h) \in [a, b]\}$ with

$$\frac{\# S(t; h)}{\{j \in \mathbb{N} : E_j(t; h) \in [a, b]\}} > 1 - \epsilon. \tag{2B.10}$$

Proof. This is a direct application of (1B.4). □

We also note that we have a global-in-time bound

$$\partial_t E_j(t;h) \leq M < \infty \tag{2B.11}$$

from differentiation of the expression

$$E_j(t;h) = \langle \mathcal{P}_h(t) u_j(t;h), u_j(t;h) \rangle \tag{2B.12}$$

and using a routine elliptic parametrix construction that is uniform in $t \in (0,1)$ to bound the quantity

$$\langle \partial_t \mathcal{P}_h(t) u_j(t;h), u_j(t;h) \rangle \tag{2B.13}$$

given that $E_j(t;h)$ lies in a fixed energy band $[a,b]$.

Recalling (2A.5), Proposition 2.8 implies that

$$\langle \partial_t \mathcal{P}_h(t) u_j, u_j \rangle \in [Q_- - \epsilon, Q_+ + \epsilon] \tag{2B.14}$$

for all $j \in S(t;h)$ such that E_j is smooth at t, and all $h < h_0(t,\epsilon)$.

Now, from the outer regularity of the Lebesgue measure, we may then choose a subinterval $J \subset (0,\delta)$ such that

$$\frac{m(\mathcal{B} \cap J)}{m(J)} > 1 - \epsilon. \tag{2B.15}$$

We can then apply the monotone convergence theorem to upgrade Proposition 2.8 for $t \in \mathcal{B}$ to a statement that is uniform in a large-measure subset of J.

Proposition 2.9. *There exists a subset $\widetilde{\mathcal{B}} \subseteq \mathcal{B} \cap J$ and an $h_0(\epsilon) > 0$ such that*

$$\frac{m(\widetilde{\mathcal{B}})}{m(J)} > 1 - 2\epsilon \tag{2B.16}$$

and, for any $h < h_0(\epsilon)$ and any $t \in \widetilde{\mathcal{B}}$, there exists a subset

$$Z(t,h) \subset \{j \in \mathbb{N} : E_j(t,h) \in [a,b]\} \tag{2B.17}$$

such that

$$\frac{\#Z(t,h)}{\#\{j \in \mathbb{N} : E_j(t,h) \in [a,b]\}} > 1 - 2\epsilon \quad \textit{for all } 0 < h < h_0 \tag{2B.18}$$

and

$$\langle \partial_t \mathcal{P}_h(t) u_j, u_j \rangle \in [Q_- - \epsilon, Q_+ + \epsilon] \quad \textit{for all } j \in Z(t,h), \tag{2B.19}$$

for all $Z(t;h)$ such that E_j is smooth at t and all $h < h_0(\epsilon)$.

Proof. For fixed $\eta, \epsilon > 0$, we define

$$\mathcal{B}_\eta := \{t \in \mathcal{B} \cap J : h_0(t,\epsilon) > \eta\}, \tag{2B.20}$$

where $h_0(t,\epsilon)$ is as in Proposition 2.8. As $\mathcal{B} \cap J = \bigcup_{\eta > 0} \mathcal{B}_\eta$, countable additivity implies that for sufficiently small $\eta_0 > 0$, we have

$$m(\mathcal{B}_{\eta_0}) > \frac{1 - 2\epsilon}{1 - \epsilon} m(\mathcal{B} \cap J) > (1 - 2\epsilon) m(J). \tag{2B.21}$$

We now take $\widetilde{\mathcal{B}} = \mathcal{B}_{\eta_0}$ and $Z(t;h) = S(t;h)$ in the notation of Proposition 2.8, and (2B.19) follows from (2B.14). □

In light of Proposition 2.9, we redefine Q_-, Q_+ to be the endpoints of the enlarged interval in (2B.19). Hence

$$\langle \partial_t \mathcal{P}_h(t) u_j, u_j \rangle \in [Q_-, Q_+] \quad \text{for all } j \in Z(t,h). \tag{2B.22}$$

In terms of the redefined Q_-, Q_+, we have

$$Q_- - B > c - \epsilon > 0, \tag{2B.23}$$

and so we have established a discrepancy between the typical speed of eigenvalue flow and the upper bound for the speed of quasieigenvalue flow.

2C. ***Spectral nonconcentration.*** We can now complete the proof of Theorem 1.1 by proving a spectral nonconcentration result that follows from the results of Section 2B.

Proposition 2.10. *Under the quantum ergodicity assumption $m(\mathcal{B}) > 0$ imposed in Section 2B, for sufficiently small $\epsilon > 0$ there exists $t_* \in \widetilde{\mathcal{B}} \subset \mathcal{B}$ such that*

$$\frac{N(t_*;h)}{\#\mathcal{M}_h(t_*)} < \frac{1}{2} \tag{2C.1}$$

for a sequence $h_j \to 0$, where

$$N(t;h) := \#\{j \in \mathbb{N} : E_j(t;h) \in W(t;h)\} \tag{2C.2}$$

is the number of exact eigenvalues lying in the union $W(t,h)$ of the quasieigenvalue windows as introduced in (2A.18).

Proof. The method of proof is by averaging in t and using Proposition 2.9 to show that most individual eigenfunctions cannot lie in $W(t,h)$ for a significant proportion of $t \in J$. We begin by defining

$$A_j(h) = \{t \in J : E_j(t;h) \in [a,b]\}, \tag{2C.3}$$

$$B_j(h) = \{t \in J : j \in Z(t;h)\}, \tag{2C.4}$$

$$C_j(h) = \{t \in J : E_j(t;h) \in W(t;h)\}. \tag{2C.5}$$

From Proposition 2.9, for each $t \in \widetilde{\mathcal{B}}$ we have

$$\sum_{j\in\mathbb{N}} 1_{B_j} \geq (1-2\epsilon) \sum_{j\in\mathbb{N}} 1_{A_j} \tag{2C.6}$$

for $h < h_0(\epsilon)$. Integrating, we obtain

$$\sum_{j\in\mathbb{N}} \int_{\widetilde{\mathcal{B}}} 1_{B_j}\,dt \geq (1-2\epsilon) \sum_{j\in\mathbb{N}} \int_{\widetilde{\mathcal{B}}} 1_{A_j}\,dt. \tag{2C.7}$$

Hence

$$\begin{aligned} \sum_{j\in\mathbb{N}} \int_J 1_{B_j}\,dt &\geq (1-2\epsilon) \sum_{j\in G(h)} \left(\int_J 1_{A_j}\,dt - \int_{J\setminus\widetilde{\mathcal{B}}} 1_{A_j}\,dt \right) \\ &\geq (1-2\epsilon) \sum_{j\in G(h)} \left(\int_J 1_{A_j}\,dt - 2\epsilon m(J) \right), \end{aligned} \tag{2C.8}$$

which can be rewritten as

$$\sum_{j\in\mathbb{N}} m(B_j) \geq (1-2\epsilon) \sum_{j\in G(h)} (m(A_j) - 2\epsilon m(J)). \tag{2C.9}$$

From the definitions (2A.19) and (2A.23), we know that $m(A_j) > 0$ only if $j \in G(h)$ and $m(A_j) = m(J)$ if $j \in \widetilde{G}(h)$. Thus we can estimate

$$\begin{aligned}\frac{1}{\#G(h)} \sum_{j\in\mathbb{N}} m(B_j) &\geq (1-2\epsilon)\left(\frac{\#\widetilde{G}(h)}{\#G(h)} - 2\epsilon\right)m(J) \\ &\geq (1-2\epsilon)(1-O(\epsilon))m(J) \\ &=: (1-\eta)m(J), \end{aligned}\tag{2C.10}$$

where $\limsup_{h\to 0} \eta(\epsilon; h) = o_\epsilon(1)$. Consequently we have

$$m(B_j) \geq (1-\eta^{1/2})m(A_j) \tag{2C.11}$$

for a subfamily $\mathcal{F}(h) \subset \widetilde{G}(h)$ with

$$\frac{\#\mathcal{F}(h)}{\#G(h)} \geq 1 - \eta^{1/2} - O(\epsilon) \tag{2C.12}$$

in the limit $h \to 0$, where we have made use of Proposition 2.4.

Taking $E(t; h) := E_j(t; h)$ for some $j \in \mathcal{F}$, the bound from the Hadamard variational formula (2B.22) yields

$$E(t_2; h) - E(t_1; h) \geq ((1-\eta^{1/2})Q_- - M\eta^{1/2})m(J), \tag{2C.13}$$

where M is the uniform bound on eigenvalue flow speed for eigenvalues in $[a, b]$ and $J = [t_1, t_2]$.

On the other hand, we now bound $E(t_2; h) - E(t_1; h)$ above. To do this, we define

$$\widetilde{E}(t; h) = E(t; h) - Bt \quad \text{and} \quad \tilde{\mu}_m(t; h) = \mu_m(t; h) - Bt,$$

where B was the upper bound in (2B.7). Then the transformed quasieigenvalue windows

$$\widetilde{W}_m(t; h) = [\tilde{\mu}_m(t; h) - h^{n+1}, \tilde{\mu}_m(t; h) + h^{n+1}]$$

are nonincreasing. From this it follows that if $\widetilde{E}(s; h) \in [\tilde{\mu}_m(s; h) - h^{n+1}, \tilde{\mu}_m(s; h) + h^{n+1}]$ and $m \in \mathcal{M}_h(s)$ for some $s \in J$, then $\widetilde{E}(s'; h) - \widetilde{E}(s; h) < 2h^{n+1}$, where s' is the final time $t \in J$ such that $m \in \mathcal{M}_h(t)$ and $\widetilde{E}(t; h) \in [\tilde{\mu}_m(t; h) - h^{n+1}, \tilde{\mu}_m + h^{n+1}]$. This implies $E(s'; h) - E(s; h) < 2h^{n+1} + B(s' - s)$.

Generalising this idea, we can cover each $C_j(h)$ with a finite union of intervals $\bigcup_k I_k$ with $I_k = [s_k, s_k']$ defined as follows:

(i) We define $s_0 := \inf\{t \in J : E(t; h) \in W(t; h)\}$, and we choose an $m(0) \in \mathcal{M}_h(s_0)$ such that $E(t; h) \in [\mu_{m(0)}(t; h) - h^{n+1}, \mu_{m(0)}(t; h) + h^{n+1}]$ and $m(0) \in \mathcal{M}_h(t)$ for all sufficiently small $t - s_0 > 0$.

(ii) We then define $s_0' := \sup\{t \in J : E(t; h) \in [\mu_{m(0)}(t; h) - h^{n+1}, \mu_{m(0)}(t; h) + h^{n+1}]\}$.

(iii) If $\{t \in J : t > s'_{k-1}$ and $E(t; h) \in W(t; h)\}$ is empty, we terminate the inductive process; otherwise we proceed inductively by defining $s_k := \inf\{t \in J : t > s'_{k-1}$ and $E(t; h) \in W(t; h)\}$ and choosing a corresponding $m(k) \in \mathcal{M}_h(s_k)$ such that $E(t; h) \in [\mu_{m(k)}(t; h) - h^{n+1}, \mu_{m(k)}(t; h) + h^{n+1}]$ and $m(k) \in \mathcal{M}_h(t)$ for all sufficiently small $t - s_{k-1} > 0$.

(iv) We then define $s'_k := \sup\{t \in J : E(t; h) \in [\mu_{m(k)}(t; h) - h^{n+1}, \mu_{m(k)}(t; h) + h^{n+1}]\}$.

From the Weyl asymptotics, this procedure must terminate after finitely many iterations.

Remark 2.11. In the case that $E(t; h)$ is still in a quasieigenvalue window for t arbitrarily close to, but greater than s'_k, we will have $s_{k+1} = s'_k$. This is the only kind of overlap possible between the intervals I_k. We also remark that the $m(k)$ are necessarily distinct, by the nature of this construction.

For each such interval $I_k = [s_k, s'_k]$, we have that $E(s'_k; h) - E(s_k; h) \le 2h^{n+1} + B(s'_k - s_k)$. As there can be at most $O(h^{-n})$ intervals I_k, we obtain

$$\sum_k E(s'_k; h) - E(s_k; h) \le B \sum_k (s'_k - s_k) + O(h). \tag{2C.14}$$

For such eigenvalues, we thus obtain the upper bound

$$\begin{aligned} E(t_2; h) - E(t_1; h) &\le \sum_k (E(s'_k; h) - E(s_k; h)) + \Big(m(J)(1 - \eta^{1/2}) - \sum_k (s'_k - s_k)\Big) Q_+ + m(J)\eta^{1/2} M \\ &\le (B - Q_+) \sum_k (s'_k - s_k) + m(J)(1 - \eta^{1/2}) Q_+ + m(J)\eta^{1/2} M + O(h) \\ &\le (B - Q_+) m(C_j) + ((1 - \eta^{1/2}) Q_+ + M\eta^{1/2}) m(J) + O(h) \end{aligned} \tag{2C.15}$$

in the limit $h \to 0$. Rearranging (2C.15) and using (2C.13), we arrive at

$$(Q_+ - B)\frac{m(C_j)}{m(J)} \le 2M\eta^{1/2} + (1 - \eta^{1/2})(Q_+ - Q_-). \tag{2C.16}$$

Hence by taking ϵ sufficiently small and then passing to sufficiently small $0 < h < h_0(\epsilon)$ we can bound $m(C_j)/m(J)$ by an arbitrarily small positive constant γ for all $j \in \mathcal{F}$. Hence we have

$$\begin{aligned} \int_J N(t; h)\, dt &\le \int_J \sum_{j \in \mathbb{N}} 1_{C_j}\, dt \le \int_J \gamma \sum_{j \in \mathcal{F}} 1_{A_j} + \#(G \setminus \mathcal{F})\, dt \\ &\le (\gamma \#\mathcal{F} + (\eta^{1/2} + O(\epsilon))\#G) m(J) \\ &\le (\gamma + \eta^{1/2} + O(\epsilon))(\#G) m(J), \end{aligned} \tag{2C.17}$$

where we used Proposition 2.4 in the final line. Fixing sufficiently small $\epsilon > 0$, for all $h < h_0(\epsilon)$ we have

$$\frac{1}{m(J)} \int_J \frac{N(t; h)}{\#\mathcal{M}_h(t)}\, dt \le \frac{1}{4}. \tag{2C.18}$$

It follows that for each such $h < h_0$, the set

$$\left\{ t \in J : \frac{N(t; h)}{\#\mathcal{M}_h(t)} \le \frac{1}{2} \right\} \tag{2C.19}$$

has measure at least $m(J)/2$. Taking a sequence $h_j \to 0$ and applying the Borel–Cantelli lemma completes the proof. $\square$

Remark 2.12. In fact, the above argument demonstrates the existence of a family of such t_* with measure bounded below by $|J|/2$; however, we shall only require a single such t_* in what follows.

We now prove an elementary spectral theory result that will show that the conclusion of Proposition 2.10 is in fact absurd, hence establishing that $m(\mathcal{B}) = 0$ and completing the proof of Theorem 1.1. We denote by U the h-dependent span of all eigenfunctions with eigenvalues in $W(t;h)$, and as in (4C.3), $\{(v_m(t;h), \mu_m(t;h))\}_{m\in\mathcal{M}_h(t)}$ denotes the family of quasimodes and associated quasieigenvalues.

Proposition 2.13. *For sufficiently small $h > 0$, the projections*

$$w_m(t_*, h) = \pi_U(v_m(t_*, h)) \tag{2C.20}$$

are linearly independent.

Proof. First, we show that the estimate from Definition 4.4 on the error of quasimodes implies that the projections $\pi_U(v_m(t_*, h))$ are large. In particular, for $m \in \mathcal{M}_h(t_*)$, we have

$$\begin{aligned}
\Big\|(\mathcal{P}_h(t_*) - \mu_m(t_*, h)) \sum_{j\in\mathbb{N}} \langle v_m(t_*, h), u_j(t_*, h)\rangle u_j\Big\|^2 &= O(h^{2\gamma+2}) \\
\implies \sum_{|E_j-\mu_m|>h^{n+1}} |E_j(t,h) - \mu_m(t,h)|^2 |\langle v_m(t_*, h), u_j(t_*, h)\rangle|^2 &= O(h^{2\gamma+2}) \\
\implies \pi_{U^\perp}(v_m(t_*, h)) &= O(h^{\gamma-n}).
\end{aligned}$$

Hence for sufficiently small h, we have

$$\|w_m\|^2 = \|\pi_U(v_m(t_*, h))\|^2 = 1 + O(h^{\gamma+1}) + O(h^{2\gamma-2n}). \tag{2C.21}$$

From the almost-orthogonality condition that our quasimodes v_m satisfy (see Definition 4.4), together with (2C.21), it follows that the w_m are almost orthogonal for distinct $m \in \mathcal{M}_h(t)$. In particular, for $m \neq k$, we have

$$\begin{aligned}
|\langle \pi_U(v_m(t_*, h)), \pi_U(v_k(t_*, h))\rangle| &\le |\langle v_m(t_*, h), v_k(t_*, h)\rangle| + |\langle \pi_{U^\perp}(v_m(t_*, h)), \pi_{U^\perp}(v_k(t, h))\rangle| \\
&= O(h^{\gamma+1}) + O(h^{2\gamma-2n}).
\end{aligned}$$

Hence

$$|\langle \pi_U(v_m(t_*, h)), \pi_U(v_k(t_*, h))\rangle - \delta_{k,m}| = O(h^{\gamma+1}) + O(h^{2\gamma-2n}) \tag{2C.22}$$

for all sufficiently small h. If we enumerate the quasimodes $v_m(t_*, h)$ by positive integers rather than $m \in \mathbb{Z}^n$, we can then form the Gram matrix $M(h) \in \mathrm{Mat}(\#\mathcal{M}_h(t_*), \mathbb{R})$, with entries given by

$$M_{ij}(h) = \langle w_i, w_j\rangle. \tag{2C.23}$$

Since

$$\|M - I\|_{HS} = (\#\mathcal{M}_h(t_*))(O(h^{\gamma+1}) + O(h^{2\gamma-2n})) = O(h^{\gamma+1-n}) + O(h^{2\gamma-3n}), \tag{2C.24}$$

we can invert $M = I + (M - I)$ as a Neumann series for sufficiently small h, provided the exponents of h are positive. This can be ensured by taking $\gamma > 3n/2$. Since M is nonsingular, we can therefore conclude that the functions in the collection

$$\{\pi_U(v_m(t_*, h)) : m \in \mathcal{M}_h(t_*)\} \tag{2C.25}$$

are linearly independent. □

We are now in a position to complete the proof of Theorem 1.1.

Completion of proof of Theorem 1.1. Having fixed $\epsilon > 0$ in Proposition 2.5, we showed in Proposition 2.10 that there exists a $t_* \in (0, \delta)$ at which we have the spectral nonconcentration result (2C.1) for a sequence $h_j \to 0$.

On the other hand, we showed in Proposition 2.13 that the projections $\pi_U(v_m(t_*, h))$ are $\#\mathcal{M}_h(t_*)$ linearly independent vectors in a vector space of dimension $\dim(U) = N(t_*, h) < \#\mathcal{M}_h(t_*)/2$. This contradiction completes the proof. □

3. Birkhoff normal form

In this section we construct a family of Birkhoff normal forms corresponding to a family of Gevrey smooth Hamiltonians $H(\theta, I; t)$, real-analytic in the parameter $t \in (-1, 1)$. The introduction of this parameter leads to only minor changes in the argument of [Popov 2004b].

We formulate the KAM theorem from [Popov 2004b] in Section 3B and outline the proof in Section 3C. We then complete the Birkhoff normal form construction following [Popov 2004b] in Section 3D.

In Section 3E, we compute the leading-order behaviour of this Birkhoff normal form as $t \to 0$, which was used in Proposition 2.5 to obtain an expression for the derivatives of the quasieigenvalues of the operator $\mathcal{P}_h(t)$ constructed in Section 4.

3A. *Notation.* We begin by introducing some notational conventions that will be used several times in this section.

Definition 3.1. For $s, r > 0$ we write

$$D_{s,r} := \{\theta \in \mathbb{C}^n/2\pi\mathbb{Z}^n : |\mathrm{Im}(\theta)| < s\} \times \{I \in \mathbb{C}^n : |I| < r\}, \tag{3A.1}$$

where $|\cdot|$ denotes the sup-norm on $\mathbb{C}^n$ induced by the 2-dimensional ℓ^∞ norm on $\mathbb{C}$.

These domains arise from considering the analytic extension of real-analytic Hamiltonians in action-angle variables. In this area it is common to bound derivatives of analytic functions using Cauchy estimates, which requires keeping track of shrinking sequences of domains.

For simplicity of nomenclature, we call an analytic function of several complex variables real-analytic if its restriction to a function of n real variables is real-valued.

As a final notational convenience, we use $|\cdot|$ to denote the ℓ^1 norm when applied to elements of $\mathbb{Z}^n$ throughout this paper, as well as the matrix norm induced by the sup norm on $\mathbb{C}^n$.

3B. *Formulation of the KAM theorem.* Let $D \subset \mathbb{R}^n$ be a bounded domain, and consider a completely integrable Hamiltonian $H^0(I) = H^0(\theta, I) : \mathbb{T}^n \times D \to \mathbb{R}$ in action-angle coordinates. To begin, we shall assume that this Hamiltonian is real-analytic.

In addition, we assume the nondegeneracy condition $\det(\partial^2 H/\partial I^2) \neq 0$. This assumption implies that the map relating the action variable I to the frequency $\omega = \nabla H^0(I)$ is locally invertible. In fact, we assume that

$$I \mapsto \nabla H^0(I) \tag{3B.1}$$

is a diffeomorphism from D to $\Omega \subset \mathbb{R}^n$. The inverse to this map is given by ∇g^0, where g^0 is the Legendre transform of H^0. The phase space $\mathbb{T}^n \times D$ is then foliated by the family of Lagrangian tori $\{\mathbb{T}^n \times \{I\} : I \in D\}$ that are invariant under Hamiltonian flow associated to H^0.

The KAM theorem asserts that small perturbations of $H^0(I)$, written as $H(\theta, I) = H^0(I) + H^1(\theta, I)$ on $\mathbb{T}^n \times D$ still possess a family of Lagrangian tori which fill up phase space up to a set of Liouville volume $o(1)$ in the size of the perturbation. More precisely, if $\Omega := \{\omega : \omega = \nabla_I H^0\}$ is the set of frequencies for the quasiperiodic flow of H^0, the frequencies satisfying

$$|\langle \omega, k\rangle| \geq \frac{\kappa}{|k|^\tau} \tag{3B.2}$$

for all nonzero $k \in \mathbb{Z}^n$ and fixed $\kappa > 0$ and $\tau > n-1$ also correspond to Lagrangian tori for the perturbed Hamiltonian H, provided $\|H - H^0\| < \epsilon(\kappa)$ in a suitable norm. Such frequencies are said to be nonresonant, and we denote the set of nonresonant frequencies by Ω_κ^*, suppressing the dependence on τ from our notation. These sets are obtained by taking the intersection of the sets

$$\left\{\omega \in \Omega : |\langle \omega, k\rangle| \geq \frac{\kappa}{|k|^\tau}\right\} \tag{3B.3}$$

over all nonzero $k \in \mathbb{Z}^n$, and hence $\bigcap_{\kappa>0} \Omega_\kappa^*$ is closed and perfect, with $\bigcup_{\kappa>0} \Omega_\kappa^*$ of full measure in Ω, as can be seen from the observation that

$$m\left(\left\{\omega \in \mathbb{R}^n : |\langle k, \omega\rangle| < \frac{\kappa}{|k|^\tau}\right\}\right) = O\left(\frac{\kappa}{|k|^{\tau+1}}\right). \tag{3B.4}$$

We work with the sets

$$\Omega_\kappa := \{\omega \in \Omega_\kappa^* : \operatorname{dist}(\omega, \partial\Omega) \geq \kappa\}, \tag{3B.5}$$

which have positive measure for sufficiently small κ. It is also convenient to introduce notation for the set of points of Lebesgue density in Ω_κ, which we denote by

$$\widetilde{\Omega}_\kappa := \left\{\omega \in \Omega : \frac{m(B(\omega, r) \cap \Omega_\kappa)}{m(B(\omega, r))} \to 1 \text{ as } r \to 0\right\}. \tag{3B.6}$$

From the Lebesgue density theorem we have that $m(\widetilde{\Omega}_\kappa) = m(\Omega_\kappa)$. We also note that a smooth function vanishing on Ω_κ is necessarily flat on $\widetilde{\Omega}_\kappa$.

The construction of the Birkhoff normal form is a consequence of Theorem 3.2, which is a version of the KAM theorem localised around the frequency ω which is taken as an independent parameter. The idea of treating ω as an independent parameter in this problem was originally due to Moser [1967]. This version is particularly useful for the Birkhoff normal form construction, as it makes it an easier task to check the regularity of the invariant tori with respect to the frequency parameter. To illustrate the setup of this theorem, we set

$$\Omega' = \{\omega \in \Omega : \mathrm{dist}(\omega, \Omega_\kappa) \le \kappa/2\}, \qquad D' = \nabla g^0(\Omega'). \tag{3B.7}$$

Taking $z_0 \in D'$, we let $I = z - z_0$ lie in a small ball of radius R about 0. That is, R is chosen such that $B_R(z_0) \subset D$. Taylor expanding gives us the expression

$$H^0(z) = H^0(z_0) + \langle \nabla_z H^0(z_0), I\rangle + \int_0^1 (1-t)\langle \nabla_z^2 H^0(z_0 + tI)I, I\rangle\, dt. \tag{3B.8}$$

We now take $\omega \in \Omega^0$ to be the corresponding frequency $\nabla H^0(z_0)$. The inverse of the frequency map is

$$\psi_0(\omega) = \nabla g^0(\omega), \tag{3B.9}$$

where g^0 is the Legendre transform of H^0. Hence we can write

$$H^0(z) = H^0(\psi_0(\omega)) + \langle \omega, I\rangle + \langle P^0(I;\omega)I, I\rangle, \tag{3B.10}$$

where P^0 is the quadratic remainder term in (3B.8). Expanding about the point $z_0 = \nabla g^0(\omega)$, we can write our perturbation H^1 locally as

$$H^1(\theta, z) = H^1(\theta, \nabla g^0(\omega) + I) = P^1(\theta, I; \omega). \tag{3B.11}$$

This leads us to consider perturbed real-analytic Hamiltonians in the form

$$H(\theta, I; \omega) = H^0(\psi_0(\omega)) + \langle \omega, I\rangle + P(\theta, I; \omega) =: N(I;\omega) + P(\theta, I;\omega), \tag{3B.12}$$

where

$$N(I;\omega) = H^0(\psi_0(\omega)) + \langle \omega, I\rangle, \tag{3B.13}$$

$$P(\theta, I; \omega) = \langle P^0(I;\omega)I, I\rangle + P^1(\theta, I; \omega). \tag{3B.14}$$

The traditional formulations of the KAM theorem assert the existence of a Cantor family of tori that persist under small perturbations of a single Hamiltonian H^0 with domain D. In the framework laid out above, we now have a Cantor family of Hamiltonians parametrised by $\omega \in \Omega_\kappa$. Note that each of these Hamiltonians consists of a component $N(I;\omega)$ that is only linear in I, and a nonlinear perturbation $P(\theta, I; \omega)$.

The essence of the frequency-localised KAM theorem in Theorem 3.2 is that for sufficiently small P we can find a symplectic change of variables that transforms H to a linear normal form in I with remainder quadratic in I for $\omega \in \Omega_\kappa$. This establishes the persistence of the Lagrangian torus with frequency ω. From Theorem 3.2, one can obtain Theorem 3.9, which establishes the existence of a Cantor family of invariant tori for the original Hamiltonian H as with traditional formulations of the KAM theorem.

To work with Gevrey smooth Hamiltonians, we fix $L_2 \ge L_0 \ge 1$ and $A_0 > 1$, and assume that $H^0 \in G^{\rho,1}_{L_0,L_2}(D^0 \times (-1,1))$ and $g^0 \in G^{\rho,1}_{L_0,L_2}(\Omega^0)$ with the estimates

$$\|H^0\|_{L_0,L_2}, \|g^0\|_{L_0,L_2} \le A_0. \tag{3B.15}$$

For $L_2 \ge L_1 \ge 1$ we now consider the analytic family of Gevrey perturbations

$$H^1 \in G^{\rho,\rho,1}_{L_1,L_2,L_2}(\mathbb{T}^n \times D \times (-1,1)),$$

with the perturbation norm

$$\epsilon_H := \kappa^{-2} \|H^1\|_{L_1,L_2,L_2}. \tag{3B.16}$$

The estimate (3B.15) implies that there is a constant $C(n,\rho)$ dependent only on n and ρ such that taking

$$R \le \frac{C(n,\rho)\kappa}{A_0 L_0^2} \tag{3B.17}$$

is sufficient to ensure that $B_R(z_0) \subset D$ for any $z_0 \in D'$.

At this point we introduce the notational convention for this section that C represents an arbitrary positive constant, dependent only on n, τ, ρ and L_0. Similarly, c will represent a positive constant strictly less than 1, also only dependent on n, τ, ρ and L_0. We will be explicit when we stray from this convention.

The estimates (3B.15) and (3B.16), together with Proposition A.3 in [Popov 2004b] show that our constructed functions P^0 and P^1 are in the Gevrey classes

$$G^{\rho}_{CL_0,CL_2,CL_2}(B_R \times \overline{\Omega}' \times (-1,1)) \subset G^{\rho}_{CL_2,CL_2}(B_R \times \overline{\Omega}' \times (-1,1)),$$
$$G^{\rho,\rho,\rho,1}_{L_1,L_2,CL_2,L_2}(\mathbb{T}^n \times B_R \times \overline{\Omega}' \times (-1,1))$$

respectively, where the C does not depend on L_0 or L_2. Additionally we have the estimate

$$\|P^1\|_{L_1,CL_2,CL_2,CL_2} \le \kappa^{-2}\epsilon_H. \tag{3B.18}$$

Dropping the factors in our Gevrey constants dependent only on n, τ, ρ, L_0 for brevity of notation, we are in a position to state the local KAM theorem in terms of the weighted norm

$$\langle P\rangle_r := r^2 \|P^0\|_{L_2,L_2,L_2} + \|P^1\|_{L_1,L_2,L_2,L_2} \tag{3B.19}$$

for $0 < r < R$.

Theorem 3.2. *Suppose $0 < \zeta \le 1$ is fixed and $\kappa < L_2^{-1-\zeta}$. Then there exists $N(n,\rho,\tau) > 0$ and $\epsilon > 0$ independent of κ, L_1, L_2, R, Ω such that whenever the Hamiltonian*

$$H(\theta, I; \omega, t) = H^0(\psi_0(\omega); t) + \langle \omega, I\rangle + \langle P^0(I;\omega,t)I, I\rangle + P^1(\theta, I; \omega, t) \tag{3B.20}$$

and $0 < r < R$ are such that

$$\langle P\rangle_r < \epsilon\kappa r L_1^{-N} \tag{3B.21}$$

we can find

$$\phi \in G^{\rho(\tau+1)+1,1}\big(\Omega \times \big(-\tfrac{3}{4}, \tfrac{3}{4}\big), \Omega\big)$$

and

$$\Phi = (U, V) \in G^{\rho,\rho(\tau+1)+1,1}\big(\mathbb{T}^n \times \Omega \times \big(-\tfrac{3}{4}, \tfrac{3}{4}\big), \mathbb{T}^n \times B_R\big)$$

such that

(i) *For all $\omega \in \Omega_\kappa$ and all $t \in \big(-\frac{3}{4}, \frac{3}{4}\big)$, the map $\Phi_{\omega,t} = \Phi(\,\cdot\,;\omega,t) : \mathbb{T}^n \to \mathbb{T}^n \times B_R$ is a G^ρ embedding, with image $\Lambda_{\omega,t}$ an invariant Lagrangian torus with respect to the Hamiltonian $H_{\phi(\omega,t),t}(\theta, I) = H(\theta, I; \phi(\omega,t), t)$. The Hamiltonian vector field restricted to this torus is given by*

$$X_{H_{\phi(\omega,t),t}} \circ \Phi_{\omega,t} = D\Phi_{\omega,t} \cdot \mathcal{L}_\omega, \tag{3B.22}$$

where

$$\mathcal{L}_\omega = \sum_{j=1}^{n} \omega_j \frac{\partial}{\partial \theta_j} \in T\mathbb{T}^n. \tag{3B.23}$$

(ii) *There exist positive constants A and C dependent only on n, τ, ρ, L_0 such that*

$$|\partial_\theta^\alpha \partial_\omega^\beta (U(\theta;\omega,t)-\theta)| + r^{-1}|\partial_\theta^\alpha \partial_\omega^\beta V(\theta;\omega,t)| + \kappa^{-1}|\partial_\omega^\beta(\phi(\omega;t)-\omega)|$$
$$\le A(CL_1)^{|\alpha|}(CL_1^{\tau+1}/\kappa)^{|\beta|}\alpha!^{\rho}\,\beta!^{\rho(\tau+1)+1}\,\frac{\langle P\rangle_r}{\kappa r}L_1^N \tag{3B.24}$$

uniformly in $\mathbb{T}^n \times \Omega \times \left(-\frac{3}{4}, \frac{3}{4}\right)$.

We remark that at the endpoint $t=0$, *this result is trivial by taking* $\phi(\omega,0)=\omega$, $U(\theta,\omega,0)=\theta$ *and* $V(\theta,\omega,0)=\nabla g^0(\omega)$.

Theorem 3.2 can be proved in the same way as [Popov 2004b, Theorem 2.1], based on the rapidly converging iterative procedure introduced in [Kolmogorov 1954]. Indeed, much of the technicality in [Popov 2004b] involves the approximation of Gevrey class Hamiltonians by real-analytic Hamiltonians. Thanks to the assumption of analyticity in t in Theorem 3.2, no such approximation is necessary in the t-parameter.

In the next section, we sketch the key steps in the proof of Theorem 3.2, highlighting the points at which the presence of the t-parameter requires a modification of the argument in [Popov 2004b].

First, we discuss the result that will comprise the steps of the iterative construction. Given a Hamiltonian in the form

$$\begin{aligned} H(\theta,I;\omega,t) &= e(\omega;t) + \langle \omega, I\rangle + P(\theta,I;\omega,t) \\ &= N(I;\omega,t) + P(\theta,I;\omega,t), \end{aligned} \tag{3B.25}$$

we aim to construct a t-dependent symplectomorphism Φ and a t-dependent frequency transformation ϕ such that for $\mathcal{F}=(\Phi,\phi)$ we have

$$(H\circ\mathcal{F})(\theta,I;\omega,t) = N_+(I;\omega,t) + P_+(\theta,I;\omega,t), \tag{3B.26}$$

where $N_+(I,\omega,t) = e_+(\omega) + \langle I,\omega\rangle$ and with $|P_+|$ controlled by $|P|^r$ for some $r>1$. This construction is analogous to that in [Pöschel 2001].

Theorem 3.3. *Suppose* $\epsilon, h, v, s, r, \eta, \sigma, K$ *are positive constants such that*

$$s, r < 1, \quad v < \tfrac{1}{6}, \quad \eta < \tfrac{1}{8}, \quad \sigma < \tfrac{1}{5}s, \quad \epsilon \le c\kappa\eta r\sigma^{\tau+1}, \quad \epsilon \le cvhr, \quad h \le \kappa/2K^{\tau+1}, \tag{3B.27}$$

where c is a constant dependent only on n and τ.

Suppose $H(\theta,I;\omega,t) = N(I;\omega,t) + P(\theta,I;\omega,t)$ *is real-analytic on* $D_{s,r}\times O_h \times (-1,1)$, *and* $|P|_{s,r,h}\le\epsilon$. *Here,* $D_{s,r}$ *is as in Definition 3.1,*

$$O_h := \{\omega \in \mathbb{C}^n : \operatorname{dist}(\omega,\Omega_\kappa) < h\}, \tag{3B.28}$$

and $|\cdot|_{s,r,h}$ *denotes the sup-norm on* $D_{s,r}\times O_h$. *Then there exists a real-analytic map*

$$\mathcal{F} = (\Phi,\phi) : D_{s-5\sigma,\eta r}\times O_{(1/2-3v)h}\times(-1,1) \to D_{s,r}\times O_h, \tag{3B.29}$$

where the maps

$$\Phi : D_{s-5\sigma,\eta r} \times O_h \times (-1,1) \to D_{s,r}, \tag{3B.30}$$

$$\phi : O_{(1/2-3\upsilon)h} \times (-1,1) \to O_h \tag{3B.31}$$

are such that

$$\begin{aligned} H \circ \mathcal{F} &= e_+(\omega, t) + \langle \omega, I \rangle + P_+(\theta, I; \omega, t) \\ &= N_+(I; \omega, t) + P_+(\theta, I; \omega, t) \end{aligned} \tag{3B.32}$$

and we have the new remainder estimate

$$|P_+|_{s-5\sigma,\eta r,(1/2-2\upsilon)h} \le C\left(\frac{\epsilon^2}{\kappa r \sigma^{\tau+1}} + (\eta^2 + K^n e^{-K\sigma})\epsilon\right). \tag{3B.33}$$

Moreover Φ *is symplectic for each* (ω, t) *and has second component affine in I. Finally, we have the uniform estimates on the change of variables*

$$|W(\Phi - \mathrm{id})|,\ |W(D\Phi - \mathrm{Id})W^{-1}| \le \frac{C\epsilon}{\kappa r \sigma^{\tau+1}}, \tag{3B.34}$$

$$|\phi - \mathrm{id}|,\ \upsilon h|D\phi - \mathrm{Id}| \le \frac{C\epsilon}{r}, \tag{3B.35}$$

where $W = \mathrm{diag}(\sigma^{-1}\mathrm{Id}, r^{-1}\mathrm{Id})$. *All estimates are uniform in the analytic parameter* $t \in (-1, 1)$.

This theorem is identical to [Popov 2004b, Proposition 3.2], with all estimates uniform in the parameter t. The proof is identical, with a detailed exposition in [Pöschel 2001]. The application of [Popov 2004b, Lemma 3.4] to obtain the frequency transformation ϕ is replaced by Proposition C.2 in our setting.

As in [Pöschel 2001; Popov 2000a], Theorem 3.3 can be used to prove the KAM theorem for real-analytic Hamiltonians $H(\theta, I; \omega, t)$. However, in order to treat the more general class of Gevrey smooth Hamiltonians $H \in G^{\rho,\rho,\rho,1}((\mathbb{T}^n \times D \times \Omega) \times (-1, 1))$, we require the approximation result Proposition 3.4.

3C. *Proof of the KAM theorem.* Following the proof of Theorem 3.2 in [Popov 2004b, Section 3], we extend the $P^j(\theta, I, \omega, t)$ to Gevrey functions

$$\widetilde{P}^j \in G^{\rho,\rho,1}_{CL_1,CL_2,CL_2}(\mathbb{T}^n \times \mathbb{R}^{2n} \times (-1,1)), \tag{3C.1}$$

where C depends only on n and ρ. We do this whilst preserving analyticity in t by making use of an adapted version of the Whitney extension theorem for anisotropic Gevrey classes, from Proposition 3.8.

We thus obtain the estimate

$$\|\widetilde{P}^j\| \le A L_1^{n+1} \|P^j\|, \tag{3C.2}$$

where A also only depends on n and ρ. We then cut off $\widetilde{P}^j$ without loss to have (I, ω) supported in $B_1 \times B_{\bar{R}} \subset \mathbb{R}^{2n}$, where $1 \ll \bar{R}$ is such that $\Omega^0 \subset B_{\bar{R}-1}$. From here, we suppress the tilde in our notation, as well as the factor C in our Gevrey constant. We require the following approximation result for functions in anisotropic Gevrey classes that plays a key role in the KAM iterative scheme.

Proposition 3.4. *Suppose $P \in G^{\rho,\rho,1}_{L_1,L_2,L_2}(\mathbb{T}^n \times \mathbb{R}^{2n} \times (-1,1))$ satisfies $\operatorname{supp}_{(I,\omega)}(P) \subset B_1 \times B_{\bar{R}}$. If u_j, w_j, v_j are positive real sequences monotonically tending to zero such that*

$$v_j L_2,\, w_j L_2 \le u_j L_1 \le 1, \quad v_0, w_0 \le L_2^{-1-\zeta}, \tag{3C.3}$$

where $1 \le L_1 \le L_2$ and $0 < \zeta \le 1$ are fixed, then we can find a sequence of real-analytic functions $P_j : U_j \to \mathbb{C}$ such that

$$|P_{j+1} - P_j|_{U_{j+1}} \le C(\bar{R}^n + 1)L_1^n \exp\bigl(-\tfrac{3}{4}(\rho-1)(2L_1 u_j)^{-1/(\rho-1)}\bigr)\|P\|, \tag{3C.4}$$

$$|P_0|_{U_0} \le C(\bar{R}^n + 1)\bigl(1 + L_1^n \exp\bigl(-\tfrac{3}{4}(\rho-1)(2L_1 u_0)^{-1/(\rho-1)}\bigr)\bigr), \tag{3C.5}$$

$$|\partial_x^\alpha (P - P_j)(\theta, I; \omega, t)| \le C(1 + \bar{R}^n)L_1^n L_2 \exp\bigl(-\tfrac{3}{4}(\rho-1)(2L_1 u_j)^{-1/(\rho-1)}\bigr) \tag{3C.6}$$

in $\mathbb{T}^n \times B_1 \times B_{\bar{R}} \times (-1,1)$ for $|\alpha| \le 1$, where

$$\begin{aligned} U_j^m := \{(\theta, I; \omega, t) \in \mathbb{C}^n/2\pi\mathbb{Z}^n \times \mathbb{C}^n \times \mathbb{C}^n \times \mathbb{C} :& \\ |\mathrm{Re}(\theta)| \le \pi,\ |\mathrm{Re}(I)| \le 2,\ |\mathrm{Re}(\omega)| \le \bar{R}+1,\ |\mathrm{Re}(t)| \le 1,& \\ |\mathrm{Im}(\theta)| \le 2u_j,\ |\mathrm{Im}(I)| \le 2v_j,\ |\mathrm{Im}(\omega_k)| \le 2w_j,\ |\mathrm{Im}(t)| \le (2L_2)^{-1}\}& \end{aligned} \tag{3C.7}$$

and

$$U_j := U_j^1, \tag{3C.8}$$

where we have identified $[-\pi, \pi]^n$ with $\mathbb{T}^n$ for simplicity of notation.

The proof of Proposition 3.4 can be found in [Popov 2004b, Section 3]. The first step is to extend P to functions $F_j : U_j^2 \to \mathbb{C}$ that are almost analytic in (θ, I, ω) and are analytic in t. The Gevrey estimate on t-derivatives of P implies that the Taylor expansions in t have radius of convergence L_2^{-1}, and so the expression

$$F_j(\theta + i\tilde{\theta}, I + i\tilde{I}, \omega + i\tilde{\omega}, t + i\tilde{t}) := \sum_{\mathcal{M}_j} \partial_\theta^\alpha \partial_I^\beta \partial_\omega^\gamma P(\theta, I; \omega, t) \frac{(i\tilde{\theta})^\alpha (i\tilde{I})^\beta (i\tilde{\omega})^\gamma (i\tilde{t})^\delta}{\alpha!\,\beta!\,\gamma!\,\delta!} \tag{3C.9}$$

is convergent on U_j^2, where the index set is as in [Popov 2004b].

The remainder of the proof in [Popov 2004b] can be followed without change. As P is analytic in t, we do not need to consider shrinking domains of analyticity as in the other variables.

The iterative scheme in [Popov 2004b, Section 3.3] can then be carried out, defining decreasing sequences of our parameters $s_j, r_j, h_j, \eta_j, \epsilon_j, \sigma_j, K_j$ such that the hypotheses of Theorem 3.3 are always satisfied, as well as decreasing sequences of the parameters u_j, v_j, w_j such that the hypotheses of the Proposition 3.4 are always satisfied. Due to the modifications made in Theorem 3.3 and Proposition 3.4 from their analogues in [Popov 2004b], all estimates are uniform in the analytic parameter $t \in (-1, 1)$.

Writing $U_j = U_j^1 \cap \{|I| < r_j\}$, where U_j^1 is defined as in Proposition 3.4, and applying Proposition 3.4 to the terms P^0, P^1 from (3B.14), we obtain sequences P_j^0, P_j^1 of real-analytic functions in U_j^1 that are good approximations to P^0 and P^1.

Setting

$$P_j(\theta, I; \omega, t) := \langle P_j^0(I; \omega, t)I, I\rangle + P_j^1(\theta, I; \omega, t), \tag{3C.10}$$

Proposition 3.4, together with the factors picked up during the Whitney extension of P^0, P^1 in (3C.2), implies the estimates

$$|P_0|_{U_0} \leq \tilde{\epsilon}_0, \tag{3C.11}$$

$$|P_j - P_{j-1}|_{U_j} \leq \tilde{\epsilon}_j, \tag{3C.12}$$

where $\tilde{\epsilon}_j$ is a positive sequence rapidly converging to zero.

Defining the Hamiltonian

$$H_j(\theta, I; \omega, t) = N_0(I; \omega) + P_j(\theta, I; \omega, t) = \langle \omega, I \rangle + P_j(\theta, I; \omega, t), \tag{3C.13}$$

which is real-analytic in U_j, one can now perform the KAM iterative scheme as in [Popov 2004b, Proposition 3.5], using the key ingredient of Theorem 3.3. For $j \geq 0$ we denote by $\mathcal{D}_j$ the class of real-analytic diffeomorphisms from $D_{j+1} \times O_{j+1} \times (-1, 1) \to D_j \times O_j$ of the form

$$\mathcal{F}(\theta, I; \omega, t) = (\Phi(\theta, I; \omega, t), \phi(\omega; t)) = (U(\theta; \omega, t), V(\theta, I; \omega, t), \phi(\omega; t)), \tag{3C.14}$$

where Φ is affine in I and canonical for fixed (ω, t). The domains are defined in terms of the parameters by $D_j = D_{s_j, r_j}$ and $O_j = O_{h_j}$.

Proposition 3.5. *Suppose P_j is real-analytic on U_j for each $j \geq 0$ and that we have the estimates*

$$|P_0|_{U_0} \leq \tilde{\epsilon}_0, \tag{3C.15}$$

$$|P_j - P_{j-1}|_{U_j} \leq \tilde{\epsilon}_j \tag{3C.16}$$

for each $j \geq 1$.

Then for each $j \geq 0$ we can find a real-analytic normal form $N_j(I; \omega, t) = e_j(\omega, t) + \langle \omega, I \rangle$ and a real-analytic map $\mathcal{F}^j$ given by

$$\mathcal{F}^{j+1} = \mathcal{F}_0 \circ \cdots \circ \mathcal{F}_j : D_{j+1} \times O_{j+1} \times (-1, 1) \to (D_0 \times O_0) \cap U_j, \tag{3C.17}$$

with the convention that the empty composition is the identity and where the $\mathcal{F}_j \in \mathcal{D}_j$ are such that

$$H_j \circ \mathcal{F}^{j+1} = N_{j+1} + R_{j+1}, \tag{3C.18}$$

$$|R_{j+1}|_{j+1} \leq \epsilon_{j+1}, \tag{3C.19}$$

$$|\overline{W}_j(\mathcal{F}_j - \mathrm{id})|_{j+1},\ |\overline{W}_j(D\mathcal{F}_j - \mathrm{Id})\overline{W}_j^{-1}| < \frac{C\epsilon_j}{r_j h_j}, \tag{3C.20}$$

$$|\overline{W}_0(\mathcal{F}^{j+1} - \mathcal{F}^j)|_{j+1} < \frac{C\epsilon_j}{r_j h_j}, \tag{3C.21}$$

where the constants C depend only on n and ρ and $\overline{W}_j = \mathrm{diag}(\sigma_j^{-1}\mathrm{Id}, r_j^{-1}\mathrm{Id}, h_j^{-1}\mathrm{Id})$.

To show that this iterative scheme converges in the Gevrey class $G^{\rho, \rho(\tau+1)+1, \rho(\tau+1)+1, 1}$ requires Gevrey estimates for the $\mathcal{S}_j := \mathcal{F}^{j+1} - \mathcal{F}^j$. To this end we introduce the domains

$$\widetilde{D}_j := \{(\theta, I) \in D_j : |\mathrm{Im}(\theta)| < s_j/2\}, \quad \widetilde{O}_j := \{\omega \in \mathbb{C}^n : \mathrm{dist}(\omega, \Omega_\kappa) < h_j/2\}. \tag{3C.22}$$

For multi-indices α, β with $|\beta| \le m$, we also introduce the following notation for the $(m-|\beta|)$-th Taylor remainder in the frequency variable, centred at ω:

$$R^m_\omega(\partial^\alpha_\theta \partial^\beta_\omega \mathcal{S}^j)(\theta, I, \omega', t) := \partial^\alpha_\theta \partial^\beta_\omega \mathcal{S}^j - \sum_{|\gamma| \le m-|\beta|} (\omega' - \omega)^\gamma \partial^\alpha_\theta \partial^{\beta+\gamma}_\omega \mathcal{S}^j(\theta, I, \omega, t)/\gamma!. \tag{3C.23}$$

We then have the following Gevrey estimates of [Popov 2004b, Lemma 3.6] uniformly in the t-parameter.

Lemma 3.6. *We have*

$$|\overline{W}_0 \partial^\alpha_\theta \partial^\beta_\omega \mathcal{S}^j(\theta, 0, \omega, t)| \le \hat{\epsilon} A C^{|\alpha|+|\beta|} L_1^{|\alpha|+|\beta|(\tau+1)+1} \kappa^{-|\beta|} \alpha!^\rho \, \beta!^{\rho'} \, E_j^{1/2} \tag{3C.24}$$

for all $(\theta, 0; \omega, t) \in \widetilde{D}_{j+1} \times \widetilde{O}_{j+1} \times (-1, 1)$, *where* $\rho' = \rho(\tau+1)+1$, *and*

$$\begin{aligned} &|\overline{W}_0(R^m_\omega \partial^\alpha_\theta \partial^\beta_\omega \mathcal{S}^j)(\theta, 0, \omega', t)| \\ &\qquad \le \hat{\epsilon} A C^{m+|\alpha|+1} L_1^{|\alpha|+(m+1)(\tau+1)+1} \kappa^{-m-1} \frac{|\omega - \omega'|^{m-|\beta|+1}}{(m-|\beta|+1)!} \alpha!^\rho \, (m+1)!^{\rho'} \, E_j^{1/2} \end{aligned} \tag{3C.25}$$

for all $\theta \in \mathbb{T}^n$, $\omega, \omega' \in \Omega_\kappa$ *and* $|\beta| \le m$, *where the constants* A, C *only depend on* n, ρ, τ, ζ.

We can now bound derivatives in t; we use the Cauchy estimate from Proposition C.1. This yields the following corollary.

Corollary 3.7. *We have*

$$|\overline{W}_0 \partial^\alpha_\theta \partial^\beta_\omega \partial^\gamma_t \mathcal{S}^j(\theta, 0; \omega, t)| \le \hat{\epsilon} A C^{|\alpha|+|\beta|+|\gamma|} L_1^{|\alpha|+|\beta|(\tau+1)+1} \kappa^{-|\beta|} \alpha!^\rho \, \beta!^{\rho'} \, \gamma! \, E_j^{1/2} \tag{3C.26}$$

for all $(\theta, 0; \omega, t) \in \widetilde{D}_{j+1} \times \widetilde{O}_{j+1} \times \left(-\frac{3}{4}, \frac{3}{4}\right)$, *where* $\rho' = \rho(\tau+1)+1$, *and*

$$\begin{aligned} &|\overline{W}_0(R^m_\omega \partial^\alpha_\theta \partial^\beta_\omega \partial^\gamma_t \mathcal{S}^j)(\theta, 0, \omega', t)| \\ &\qquad \le \hat{\epsilon} A C^{m+|\alpha|+|\gamma|+1} L_1^{|\alpha|+(m+1)(\tau+1)+1} \kappa^{-m-1} \frac{|\omega - \omega'|^{m-|\beta|+1}}{(m-|\beta|+1)!} \alpha!^\rho \, (m+1)!^{\rho'} \, \gamma! \, E_j^{1/2} \end{aligned} \tag{3C.27}$$

for all $\theta \in \mathbb{T}^n$, $\omega, \omega' \in \Omega_\kappa$, $t \in \left(-\frac{3}{4}, \frac{3}{4}\right)$ *and* $|\beta| \le m$, *where the constants* A, C *only depend on* n, ρ, τ, ζ.

From Proposition 3.5 and Corollary 3.7, the rapid decay of E_j implies that the limit

$$\partial^\alpha_\theta \partial^\gamma_t \mathcal{H}^\beta(\theta, \omega; t) := \lim_{j \to \infty} \partial^\alpha_\theta \partial^\beta_\omega \partial^\gamma_t (\mathcal{F}^j(\theta, 0; \omega, t) - (\theta, 0, \omega)) \tag{3C.28}$$

exists for each $(\theta; \omega, t) \in \mathbb{T}^n \times \Omega_\kappa \times \left(-\frac{3}{4}, \frac{3}{4}\right)$, and each triple of multi-indices α, β, γ. Convergence is uniform, and the limit is smooth in θ and t and continuous in ω, with $\partial^\alpha_\theta \partial^\gamma_t (\mathcal{H}^\beta) = \partial^\alpha_\theta \partial^\gamma_t \mathcal{H}^\beta$, justifying the notation in (3C.28).

We now need to use the jet $\mathcal{H} = (\partial^\alpha_\theta \partial^\gamma_t \mathcal{H}^\beta)$ of continuous functions $\mathbb{T}^n \times \Omega_\kappa \times \left(-\frac{3}{4}, \frac{3}{4}\right) \to \mathbb{T}^n \times D \times \Omega$ to obtain a Gevrey function on $\mathbb{T}^n \times \Omega \times \left(-\frac{3}{4}, \frac{3}{4}\right)$ by using a Gevrey version of the Whitney extension theorem. We define

$$(R^m_\omega \partial^\alpha_\theta \partial^\gamma_t \mathcal{H})_\beta(\theta, \omega', t) := \partial^\alpha_\theta \partial^\gamma_t \mathcal{H}^\beta(\theta, \omega', t) - \sum_{|\delta| \le m-|\beta|} (\omega' - \omega)^\delta \partial^\alpha_\theta \partial^\gamma_t \mathcal{H}^{\beta+\delta}(\theta; \omega, t)/\gamma!. \tag{3C.29}$$

In this notation, the results of Corollary 3.7 yield

$$|\overline{W}_0\partial_\theta^\alpha\partial_t^\gamma\mathcal{H}^\beta(\theta;\omega,t)|\le\hat{\epsilon}AL_1(CL_1)^{|\alpha|}(CL_1^{\tau+1}/\kappa)^{|\beta|}C^\gamma\alpha!^\rho\,\beta!^{\rho'}\,\gamma!,\tag{3C.30}$$

$$|\overline{W}_0(R_\omega^m\partial_\theta^\alpha\partial_t^\gamma\mathcal{H})_\beta(\theta,\omega',t)|\le\hat{\epsilon}AL_1(CL_1)^{|\alpha|}(CL_1^{\tau+1}/\kappa)^{m+1}C^\gamma\frac{|\omega-\omega'|^{m-|\beta|+1}}{(m-|\beta|+1)!}\alpha!^\rho\,(m+1)!^{\rho'}\,\gamma!\tag{3C.31}$$

for $|\beta|\le m$, and $(\theta,\omega,\omega',t)\in\mathbb{T}^n\times\Omega_\kappa\times\Omega_\kappa\times\left(-\frac{3}{4},\frac{3}{4}\right)$, where A and C depend only on n,ρ,τ. These estimates allow us to apply the following consequence of Theorem D.3.

Proposition 3.8. *Suppose $K\subset\mathbb{R}^n$ is compact, and $1\le\rho<\rho'$. If the jet $(f^{\alpha,\beta,\gamma})$ of functions $f^{\alpha,\beta,\gamma}:\mathbb{T}^n\times K\times\left(-\frac{3}{4},\frac{3}{4}\right)\to\mathbb{R}$ is continuous on $\mathbb{T}^n\times K\times\left(-\frac{3}{4},\frac{3}{4}\right)$ and is smooth in $(\theta,t)\in\mathbb{T}^n\times\left(-\frac{3}{4},\frac{3}{4}\right)$ for each fixed $\omega\in K$, where*

$$\partial_\theta^{\alpha'}\partial_t^{\gamma'}(f^{\alpha,\beta,\gamma})=f^{\alpha+\alpha',\beta,\gamma+\gamma'},\tag{3C.32}$$

and we have the estimates

$$|f^{\alpha,\beta,\gamma}(\theta;\omega,t)|\le AC_1^{|\alpha|}C_2^{|\beta|}C_3^{|\gamma|}\alpha!^\rho\,\beta!^{\rho'}\,\gamma!,\tag{3C.33}$$

$$|(R_\omega^m\partial_\theta^\alpha\partial_t^\gamma f)_\beta(\theta,\omega',t)|\le AC_1^{|\alpha|}C_2^{m+1}C_3^{|\gamma|}\frac{|\omega-\omega'|^{m-|\beta|+1}}{(m-|\beta|+1)!}\alpha!^\rho\,(m+1)!^{\rho'}\,\gamma!\tag{3C.34}$$

then there exist positive constants A_0, C_0, dependent only on (n,ρ,τ) (in particular, independent of the set K) such that we can extend f to $\tilde{f}\in G^{\rho,\rho',1}\left(\mathbb{T}^n\times\mathbb{R}^n\times\left(-\frac{3}{4},\frac{3}{4}\right)\right)$ such that $\partial_\theta^\alpha\partial_\omega^\beta\partial_t^\gamma\tilde{f}=f^{\alpha,\beta,\omega}$ on $\mathbb{T}^n\times K\times\left(-\frac{3}{4},\frac{3}{4}\right)$ and

$$|\partial_\theta^\alpha\partial_\omega^\beta\partial_t^\gamma\tilde{f}(\theta,\omega)|\le A_0A\max(C_1,1)C_0^{|\alpha|+|\beta|+|\gamma|+n}C_1^{|\alpha|+n}C_2^{|\beta|}C_3^{|\gamma|}\alpha!^\rho\,\beta!^{\rho'}\,\gamma!\,.\tag{3C.35}$$

The proof of Proposition 3.8 is identical to that in [Popov 2004b, Theorem 3.7], making use of Theorem D.3 involving the parameter t. Having established Proposition 3.8, the proof of Theorem 3.2 can be completed as in [Popov 2004b, Section 3.5] without modification.

3D. *Birkhoff normal form.* We obtain a Birkhoff normal form for near-integrable Hamiltonians using a version of the KAM theorem that is a consequence of Theorem 3.2. The Gevrey index $\rho(\tau+1)+1$ frequently appears in these results, and so we introduce $\rho':=\rho(\tau+1)+1$.

Theorem 3.9. *Fix $0<\zeta\le1$ and let $H^0(I;t)$ be a real-valued nondegenerate smooth family of Hamiltonians in $G^{\rho,1}(D^0\times(-1,1))$ and let D be a subdomain with $\overline{D}\subset D^0$. We define $\Omega=\nabla H^0(D)$ and fix $L_2\ge L_1\ge1$ and $\kappa\le L_2^{-1-\zeta}$ such that $L_2\ge L_0$ and $\Omega_\kappa\ne\varnothing$. Then there exists $N=N(n,\rho,\tau)$ and $\epsilon>0$ independent of κ, L_1, L_2 and $D\subset D^0$ such that for any $H\in G^{\rho,\rho,1}_{L_1,L_2,L_2}(\mathbb{T}^n\times D\times(-1,1))$ with norm*

$$\epsilon_H:=\kappa^{-2}\|H-H^0\|_{L_1,L_2,L_2}\le\epsilon L_1^{-N}\tag{3D.1}$$

there exists a map

$$\overline{\Phi}=(\overline{U},\overline{V})\in G^{\rho,\rho',1}\left(\mathbb{T}^n\times\Omega\times\left(-\tfrac{3}{4},\tfrac{3}{4}\right),\mathbb{T}^n\times D\right)\tag{3D.2}$$

such that:

(i) *For each $\omega\in\Omega_\kappa$ and each $t\in\left(-\frac{3}{4},\frac{3}{4}\right)$, $\Lambda_\omega=\{\overline{\Phi}(\theta;\omega,t):\theta\in\mathbb{T}^n\}$ is an embedded invariant Lagrangian torus of H, and $X_H\circ\overline{\Phi}(\,\cdot\,;\omega,t)=D\overline{\Phi}(\,\cdot\,;\omega,t)\cdot\mathcal{L}_\omega$.*

(ii) *There exist constants $A, C > 0$ independent of κ, L_1, L_2 and $D \subset D^0$ such that*

$$|\partial_\theta^\alpha \partial_\omega^\beta (\overline{U}(\theta;\omega,t)-\theta)| + \kappa^{-1}|\partial_\theta^\alpha \partial_\omega^\beta (\overline{V}(\theta;\omega,t)-\nabla g^0(\omega))| \\ \le A(CL_1)^{|\alpha|}(CL_1^{\tau+1}/\kappa)^{|\beta|}\alpha!^{\rho}\,\beta!^{\rho'}\,L_1^{N/2}\epsilon_H^{1/2} \quad \text{(3D.3)}$$

uniformly in $\mathbb{T}^n \times \Omega \times \left(-\frac{3}{4}, \frac{3}{4}\right)$.

The proof of Theorem 3.9 is identical to [Popov 2004b, Theorem 1.1], making use of Theorem 3.2. We can now use Theorem 3.9 to obtain the Birkhoff normal form as done in [Popov 2004b].

Theorem 3.10. *Suppose the assumptions of Theorem 3.9 hold. Then there exist $N(n,\rho,\tau) > 0$ and $\epsilon > 0$ independent of κ, L_1, L_2, D such that for any $H \in G^{\rho,\rho,1}_{L_1,L_2,L_2}(\mathbb{T}^n \times D \times (-1,1))$ with*

$$\epsilon_H \le \epsilon L_1^{-N-2(\tau+2)}, \quad \text{(3D.4)}$$

where ϵ_H is as in (3D.1), there is a family of $G^{\rho',\rho'}$ maps $\omega : D \times \left(-\frac{1}{2},1,2\right) \to \Omega$ and a family of maps $\chi \in G^{\rho,\rho',\rho'}\left(\mathbb{T}^n \times D \times \left(-\frac{1}{2},1,2\right), \mathbb{T}^n \times D\right)$ that are diffeomorphisms and exact symplectomorphisms respectively for each fixed $t \in \left(-\frac{1}{2},1,2\right)$. Moreover, we can choose the maps ω and χ such that family of transformed Hamiltonians

$$\widetilde{H}(\theta, I; t) := (H \circ \chi)(\theta, I; t) \quad \text{(3D.5)}$$

is of Gevrey class $G^{\rho,\rho',\rho'}\left(\mathbb{T}^n \times D \times \left(-\frac{1}{2},1,2\right)\right)$ and can be decomposed as

$$K(I;t) + R(\theta, I; t) := \widetilde{H}(0, I; t) + (\widetilde{H}(\theta, I; t) - \widetilde{H}(0, I; t)) \quad \text{(3D.6)}$$

such that:

(i) *$\mathbb{T}^n \times \{I\}$ is an invariant Lagrangian torus of $\widetilde{H}(\cdot,\cdot;t)$ for each $I \in E_\kappa(t) = \omega^{-1}(\widetilde{\Omega}_\kappa; t)$ and each $t \in \left(-\frac{1}{2},1,2\right)$.*

(ii) *$\partial_I^\beta(\nabla K(I;t) - \omega(I;t)) = \partial_I^\beta R(\theta, I; t) = 0$ for all $(\theta, I; t) \in \mathbb{T}^n \times E_\kappa(t) \times \left(-\frac{1}{2},1,2\right)$, $\beta \in \mathbb{N}^n$.*

(iii) *There exist $A, C > 0$ independent of κ, L_1, L_2, and $D \subset D^0$ such that we have the estimates*

$$|\partial_\theta^\alpha \partial_I^\beta \partial_t^\delta \phi(\theta, I; t)| + |\partial_I^\beta \partial_t^\delta(\omega(I;t) - \nabla H^0(I;t))| + |\partial_\theta^\alpha \partial_I^\beta \partial_t^\delta(\widetilde{H}(\theta, I; t) - H^0(I;t))| \\ \le A\kappa C^{|\alpha|+|\beta|+|\delta|} L_1^{|\alpha|}(L_1^{\tau+1}/\kappa)^{|\beta|}\alpha!^{\rho}\,\beta!^{\rho'}\,\delta!^{\rho'}\,L_1^{N/2}\epsilon_H^{1/2} \quad \text{(3D.7)}$$

uniformly in $\mathbb{T}^n \times D \times \left(-\frac{1}{2},1,2\right)$ for all α, β, where $\phi \in G^{\rho,\rho',\rho'}\left(\mathbb{T}^n \times D \times \left(-\frac{1}{2},1,2\right)\right)$ is such that $\langle \theta, I\rangle + \phi(\theta, I; t)$ generates the symplectomorphism χ in the sense of (3E.8).

Remark 3.11. For our purposes, high regularity in the t-parameter is not required, so we have dropped from analyticity to $G^{\rho'}$ regularity in t at this point in order to simplify the proceeding arguments. We expect that analyticity in t could be preserved by using a stronger variant of the Komatsu implicit function theorem than Corollary A.5.

Proof. We begin by taking ϵ, N as in Theorem 3.9 and noting that $\epsilon_H \le \epsilon L_1^{-N-2}$ by assumption. This implies that the factor $(ACL_1)L_1^{N/2}\sqrt{\epsilon_H}$ occurring in the Gevrey estimate (3D.3) can be bounded above by $AC\sqrt{\epsilon}$. Hence, taking ϵ small enough that both the conclusion to Theorem 3.9 holds as well as $AC\sqrt{\epsilon} < \frac{1}{2}$, we can first apply the Cauchy estimate from Proposition C.1 to (3D.3) in t, and then apply a variant of the Komatsu implicit function theorem, Corollary A.5, to obtain a solution $\theta(\gamma;\omega,t):\mathbb{T}^n\times\Omega\times\left(-\frac{1}{2},1,2\right)\to\mathbb{T}^n$ to the implicit equation

$$\bar{U}(\theta;\omega,t)=\gamma. \tag{3D.8}$$

Moreover, this solution satisfies the Gevrey estimate

$$|\partial_\gamma^\alpha\partial_\omega^\beta\partial_t^\delta(\theta(\gamma;\omega,t)-\gamma)|\le AC^{|\alpha|+|\beta|+|\delta|}L_1^{|\alpha|}(L_1^{\tau+1}/\kappa)^{|\beta|}\alpha!^\rho\,\beta!^{\rho'}\delta!^{\rho'}\,L_1^{N/2}\sqrt{\epsilon_H} \tag{3D.9}$$

uniformly on $\mathbb{T}^n\times\Omega\times\left(-\frac{1}{2},1,2\right)$.

We set $F(\gamma;\omega,t):=\bar{V}(\theta(\gamma;\omega,t);\omega,t)$. In terms of $(\gamma;\omega,t)$, the Lagrangian torus Λ_ω is now given by $(\gamma,F(\gamma;\omega,t):\gamma\in\mathbb{T}^n)$ for each $\omega\in\Omega_\kappa$ and each $t\in\left(-\frac{1}{2},1,2\right)$. Moreover, Proposition A.7 on the composition of Gevrey functions gives us the estimate

$$|\partial_\gamma^\alpha\partial_\omega^\beta\partial_t^\delta(F(\gamma;\omega,t)-\nabla g^0(\omega))|\le A\kappa C^{|\alpha|+|\beta|+|\delta|}L_1^{|\alpha|}(L_1^{\tau+1}/\kappa)^{|\beta|}\alpha!^\rho\,\beta!^{\rho'}\,\delta!^{\rho'}\,L_1^{N/2}\sqrt{\epsilon_H}. \tag{3D.10}$$

We next construct functions $\psi\in G^{\rho,\rho',\rho'}\left(\mathbb{R}^n\times\Omega\times\left(\frac{1}{2},1,2\right)\right)$ and $R\in G^{\rho',\rho'}\left(\Omega\times\left(\frac{1}{2},1,2\right)\right)$ such that the function

$$Q(x;\omega,t):=\psi(x;\omega,t)-\langle x,R(\omega,t)\rangle \tag{3D.11}$$

is 2π-periodic in x and satisfies

$$\nabla_x\psi(x;\omega,t)=F(p(x),\omega,t) \tag{3D.12}$$

in $\mathbb{R}^n\times\Omega_\kappa\times\left(\frac{1}{2},1,2\right)$, where $p:\mathbb{R}^n\to\mathbb{T}^n$ is the canonical projection, as well as the estimate

$$\begin{aligned}|\partial_x^\alpha\partial_\omega^\beta\partial_t^\delta Q(x;\omega,t)|&+|\partial_\omega^\beta\partial_t^\delta(R(\omega,t)-\nabla g^0(\omega))|\\&\le A\kappa C^{|\alpha|+|\beta|+|\delta|}L_1^{|\alpha|}(L_1^{\tau+1}/\kappa)^{|\beta|}\alpha!^\rho\,\beta!^{\rho'}\,\delta!^{\rho'}\,L_1^{N/2}\sqrt{\epsilon_H}\end{aligned} \tag{3D.13}$$

for $(x;\omega,t)\in\mathbb{R}^n\times\Omega\times\left(\frac{1}{2},1,2\right)$.

We do this by first integrating the canonical 1-form $I\,dx$ over the chain

$$c_x:=\{(sx,F(p(sx);\omega,t)):0\le s\le 1\}\subset\mathbb{R}^n\times D. \tag{3D.14}$$

We define

$$\tilde{\psi}(x;\omega,t):=\int_{c_x}\sigma=\int_0^1\langle F(p(sx);\omega,t),x\rangle\,ds \tag{3D.15}$$

in $\mathbb{R}^n\times\Omega\times\left(\frac{1}{2},1,2\right)$. From the estimate (3D.10) it follows that $\tilde{\psi}(x;\omega,t)-\langle\nabla g^0(\omega),x\rangle$ is bounded above by the right-hand side of (3D.13) in $[0,4\pi]^n\times\Omega\times\left(\frac{1}{2},1,2\right)$. Hence if we define $R_j(\omega,t)=(2\pi)^{-1}\tilde{\psi}(2\pi e_j;\omega,t)$, then $R-\nabla g^0$ satisfies the required estimates in (3D.13).

Since for $\omega\in\Omega_\kappa$ we know that Λ_ω is a Lagrangian torus, it follows that the integral of the canonical 1-form over any closed chain in Λ_ω is homotopy invariant. This means that such an integral is a homomorphism from the fundamental group of Λ_ω to $\mathbb{R}$. Hence

$$\tilde{\psi}(x+2\pi m;\omega,t)-\tilde{\psi}(x;\omega,t)=\langle 2\pi m,R(\omega,t)\rangle \tag{3D.16}$$

and so the function

$$\widetilde{Q}(x;\omega,t) := \tilde{\psi}(x,\omega) - \langle x, R(\omega,t)\rangle \tag{3D.17}$$

both satisfies the Gevrey estimate in (3D.13) and is 2π-periodic in x for $(\omega,t)\in\Omega_\kappa\times\left(\frac{1}{2},1,2\right)$.

To obtain the sought Q in (3D.11) from $\widetilde{Q}$, we use an averaging trick. Choosing $f\in G^\rho_C(\mathbb{R}^n)$ for some positive constant C such that f is supported in $[\pi/2, 7\pi/2]^n$ and

$$\sum_{k\in\mathbb{Z}^n} f(x+2\pi k) = 1 \tag{3D.18}$$

for each $x\in\mathbb{R}^n$, it then follows that

$$Q(x;\omega,t) := \sum_{k\in\mathbb{Z}^n} f(x+2\pi k)\widetilde{Q}(x+2\pi k;\omega,t) \tag{3D.19}$$

is 2π-periodic in x for every $\omega\in\Omega$ and coincides with $\widetilde{Q}$ for $\omega\in\Omega_\kappa$. Moreover, Q satisfies the same Gevrey estimate (3D.13) as $\widetilde{Q}$. We define

$$\psi(x;\omega,t) := Q(x;\omega,t) + \langle x, R(\omega,t)\rangle. \tag{3D.20}$$

Note that by multiplying Q and $R-\nabla g^0$ by a cut-off function $h\in G^{\rho'}_{C/\kappa}$ which is equal to 1 in a ω-neighbourhood of Ω_κ and vanishes for $\operatorname{dist}(\omega,\mathbb{R}^n\setminus\Omega)\le\kappa/2$, where $C>0$ is independent of $\Omega\subset\Omega^0$, we can assume that $\psi(x;\omega,t)=\langle x,\nabla g^0(\omega)\rangle$ for $\operatorname{dist}(\omega,\mathbb{R}^n\setminus\Omega)\le\kappa/2$. This cutoff preserves the Gevrey estimates on ψ.

Now since $\epsilon_H L_1^{N+2(\tau+2)}\le\epsilon$, we have that $\kappa A(CL_1)(CL_1^{\tau+1}/\kappa)L_1^{N/2}\sqrt{\epsilon_H}\le AC^2\sqrt{\epsilon}$. By taking ϵ sufficiently small we have that $\omega\mapsto\nabla_x\psi(x;\omega,t)$ is a diffeomorphism for any fixed $x\in\mathbb{R}^n$ from the Gevrey estimate (3D.13). Hence we have a $G^{\rho,\rho'}$-foliation of $\mathbb{T}^n\times D$ by Lagrangian tori $\Lambda_\omega=\{(p(x),\nabla_x\psi(x,\omega)) : x\in\mathbb{R}^n\}$, where $\omega\in\Omega$.

In the sought coordinate change, the action $I(\omega,t)$ of the Lagrangian torus Λ_ω will be given by $R(\omega,t)$. Hence from (3D.13) and Proposition A.4, it follows that for ϵ sufficiently small, the map

$$(\omega,t)\mapsto(I(\omega,t),t) = (R(\omega,t),t) \tag{3D.21}$$

is a $G^{\rho',\rho'}$-diffeomorphism and we have the Gevrey estimate

$$|\partial_I^\alpha\partial_t^\beta(\omega(I,t)-\nabla H^0(I;t))| \le A\kappa C^{|\alpha|+|\beta|}(L_1^{\tau+1}/\kappa)^{|\alpha|}\alpha!^{\rho'}\,\beta!^{\rho'}\,L_1^{N/2}\sqrt{\epsilon_H} \tag{3D.22}$$

uniformly for $(\theta,I,t)\in\mathbb{T}^n\times D\times\left(\frac{1}{2},1,2\right)$.

We construct the sought symplectomorphism χ using the generating function $\Phi(x,I;t)$, setting

$$\Phi(x,I;t) = \psi(x,\omega(I;t);t) \tag{3D.23}$$

and noting that we have the required 2π-periodicity of $\phi(x,I;t) := \Phi(x,I,t) - \langle x,I\rangle$, and from Proposition A.7, we also have the estimate

$$|\partial_x^\alpha\partial_I^\beta\partial_t^\delta(\Phi(x,I;t-\langle x,I\rangle))| \le A\kappa C^{|\alpha|+|\beta|+|\delta|}L_1^{|\alpha|}(L_1^{\tau+1}/\kappa)^{|\beta|}\alpha!^\rho\,\beta!^{\rho'}\,\delta!^{\rho'}\,L_1^{N/2}\sqrt{\epsilon_H}. \tag{3D.24}$$

We can then apply Corollary A.5 to solve the implicit equation

$$\partial_I \Phi(\gamma, I, t) = \theta \tag{3D.25}$$

for γ with the estimate

$$|\partial_\theta^\alpha \partial_I^\beta \partial_t^\delta (\gamma(\theta, I, t) - \theta)| \leq A\kappa C^{|\alpha|+|\beta|+|\delta|} L_1^{|\alpha|} (L_1^{\tau+1}/\kappa)^{|\beta|} \alpha!^\rho \beta!^{\rho'} \delta!^{\rho'} L_1^{N/2} \sqrt{\epsilon_H}. \tag{3D.26}$$

This completes the construction of a symplectomorphism χ satisfying

$$\chi(\partial_I \Phi(\theta, I, t), I) = (\theta, \partial_\theta \Phi(\theta, I, t)). \tag{3D.27}$$

It follows that

$$(\theta, F(\theta; \omega, t)) = \chi(\partial_I \Phi(\theta, I(\omega), t), I(\omega)) = \chi(\theta, I(\omega), t) \tag{3D.28}$$

for $\omega \in \Omega_\kappa$ and so

$$\Lambda_\omega = \{\chi(\theta, I(\omega), t) : \theta \in \mathbb{T}^n\} \tag{3D.29}$$

for $(\omega, t) \in \Omega_\kappa \times \left(\frac{1}{2}, 1, 2\right)$.

We now set $\widetilde{H}$, K, R as in the theorem statement in terms of the symplectomorphism χ. Since H is constant on Λ_ω for each $\omega \in \Omega_\kappa$, it follows that $R(\,\cdot\,, I; t)$ is identically zero for each $I = I(\omega)$ with $\omega \in \Omega_\kappa$. Hence R is flat at $I \in E_\kappa(t)$, since each point in $E_\kappa(t)$ is of positive density in $I(\Omega_\kappa)$.

Finally, the Gevrey estimate in (3D.7) for $\widetilde{H}(\theta, I, t) - H(I, t)$ follows from Proposition A.7. □

3E. *Calculation of $\partial_t K_0(I, 0)$*. A crucial ingredient in the proof of Theorem 1.1 is the calculation of the derivative of quasieigenvalues in Proposition 2.5 in the semiclassical limit $h \to 0$. From the truncated quantum Birkhoff normal form in Theorem 4.1, this can be reduced to the study of the t-dependence of the integrable term $K(I; t)$ in the classical Birkhoff normal form established in Theorem 3.10.

We now consider a 1-parameter family of Hamiltonians $H(\theta, I; t)$ satisfying the assumptions of Theorem 1.1. We can write

$$H(\theta, I; t) = H^0(I) + H^1(\theta, I; t), \tag{3E.1}$$

with

$$H^0(I) := H(\theta, I; 0) \tag{3E.2}$$

$$H^1(\theta, I; t) := t\partial_t H(\theta, I; 0) + \int_0^t (1-s)\partial_t^2 H(\theta, I; s)\, ds = t\partial_t H(\theta, I; 0) + O(t^2), \tag{3E.3}$$

and we assume that H additionally satisfies the assumptions of Theorem 3.10 with this choice of H^0, H^1. By applying two KAM stem iterations to $H(\theta, I; t)$, we obtain a transformed completely integrable component and reduce the order of magnitude of the θ-dependent remainder. An application of Theorem 3.10 to this transformed Hamiltonian produces a Birkhoff normal form, and (3D.7) yields an expression for $K(I; t)$ up to order $o(t)$.

The KAM step iterations required differ from that in Theorem 3.3, in that they are not parametrised by $\omega \in \Omega$ and instead take place in the action-angle space $\mathbb{T}^n \times D$. Such a KAM step appears in the proof of the KAM theorem found in [Gallavotti 1983]. We first describe the KAM step without the presence of the parameter t for simplicity. One begins with a perturbation

$$H(\theta, I) = H^0(I) + H^1(\theta, I) \tag{3E.4}$$

of a completely integrable Hamiltonian $H^0(I)$, and a fixed perturbation $H^1(\theta, I)$, both analytic on the complex domain

$$\theta \in 2\pi\mathbb{C}^n \setminus 2\pi\mathbb{R}^n, \quad |\mathrm{Im}(\theta)| < s, \tag{3E.5}$$

$$\mathrm{Re}(I) \in D, \qquad |\mathrm{Re}(I)| < r. \tag{3E.6}$$

We assume that $|H^1|_{s,r} = O(\epsilon)$ in the uniform sense.

By consideration of the linearised Hamilton–Jacobi equation, we choose a symplectic transformation $\chi : \mathbb{T}^n \times D \to \mathbb{T}^n \times D$ with the aim to write

$$\widetilde{H}(\theta, I) = (H \circ \chi)(\theta, I) = \widetilde{H}^0(I) + \widetilde{H}^1(\theta, I), \tag{3E.7}$$

with $\widetilde{H}^1 = O(\epsilon^\alpha)$ for some $\alpha > 1$. Then we have transformed a sufficiently small perturbation of an integrable Hamiltonian to an even smaller perturbation of a new integrable Hamiltonian, in a way we can hope to iterate.

Obtaining the "new" error bound for $\widetilde{H}^1$ necessarily requires a shrinking of the domains of analyticity, through the use of Cauchy estimates to control derivatives. Moreover, there is a more subtle shrinking of domain required in the I-variable, due to the infamous "small-divisor" problem. Specifically, χ is found using terms of the generating function

$$\Phi(I', \theta) = i \sum_{k \in \mathbb{Z}^n : 0 < |k| \le M} \frac{H_k^1(I') e^{ik\cdot\theta}}{\omega(I') \cdot k}, \tag{3E.8}$$

where H_k^1 denotes the k-th Fourier coefficient of H^1, and $\omega = \nabla_I H^0(I)$; see [Gallavotti 1983, (2.10)].

The denominators in (3E.8) can generally be zero, and so one must restrict to values of I' for which we have a nonresonance condition

$$\omega(I') \cdot k \ge \frac{C}{|k|^2} \tag{3E.9}$$

for all $0 < |k| \le M$, where C and M are chosen suitably. We also need to remove those actions I' with $\mathrm{dist}(I', \partial\Omega) \le \tilde{\rho}$ so that the perturbed tori do not escape the coordinate patch; see [Gallavotti 1983, (3.12)] for the choice of the constant $\tilde{\rho}$. This leads to the definition of the set

$$\widetilde{D}_1 = \{I \in D : \mathrm{dist}(I, \partial D) > \tilde{\rho} \text{ and } \omega(I) \cdot k \ge C/|k|^2 \text{ for all } 0 < |k| \le M\}. \tag{3E.10}$$

For any $\tilde{I} \in \widetilde{D}_1$ the expression (3E.8) is certainly defined, but as the domain might have rather rough boundary, it is convenient to slightly enlarge $\widetilde{D}_1$ to the open set

$$D_1 = \bigcup_{I \in \widetilde{D}_1} B(I, \tilde{\rho}/2). \tag{3E.11}$$

Upon restricting to this action set for suitable C and M, the objective of (3E.7) can indeed be achieved, and the "integrable part" of the new Hamiltonian can be written as

$$\widetilde{H}^0(I) = H^0(I) + (2\pi)^{-n} \int H^1(\theta, I)\, d\theta; \tag{3E.12}$$

see [Gallavotti 1983, (3.38)]. The overall transformed Hamiltonian is then given by

$$\widetilde{H}(\tilde{\theta}, \tilde{I}) = \widetilde{H}^0(\tilde{I}) + \widetilde{H}^1(\tilde{\theta}, \tilde{I}) \tag{3E.13}$$

in the domain $\mathbb{T}^n \times D_1$, with

$$\|\widetilde{H}^1\| = O(\epsilon^{3/2}). \tag{3E.14}$$

The classical KAM theorem is then proven in [Gallavotti 1983] by iterating this procedure, carefully choosing the C, M, $\tilde{\rho}$ and the analyticity parameters r, s so that the estimate (3E.14) is satisfied with every step, ensuring convergence, and so that the limiting domain $\bigcap_j D_j$ of nonresonant actions is of large measure. A full discussion of this procedure can be found in [Gallavotti 1983].

We now return to our setting of the one-parameter family of Hamiltonians

$$H(\theta, I; t) = H^0(I) + H^1(\theta, I; t).$$

One iteration of the KAM step outlined above yields a family of symplectomorphisms

$$\chi_1 : \mathbb{T}^n \times D_1 \to \mathbb{T}^n \times D \tag{3E.15}$$

parametrised by t such that

$$\widetilde{H}(\theta, I; t) = (H \circ \chi_1)(\theta, I; t) = H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta + \widetilde{H}^1(\theta, I; t), \tag{3E.16}$$

where the second term comes from (3E.3) and the error term $\widetilde{H}^1(\theta, I; t) = O(t^{3/2})$. Regarding this transformed Hamiltonian as a small perturbation of the integrable Hamiltonian

$$\widetilde{H}^0(I; t) = H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta, \tag{3E.17}$$

we perform one more KAM iteration to obtain another family of symplectomorphisms

$$\chi_2 : \mathbb{T}^n \times D_3 \to \mathbb{T}^n \times D_2 \tag{3E.18}$$

parametrised by t such that

$$\begin{aligned}\widetilde{\widetilde{H}}(\theta, I; t) &= (\widetilde{H} \circ \chi_2)(\theta, I; t) \\ &= H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta + (2\pi)^{-n} \int_{\mathbb{T}^n} \widetilde{H}^1(\theta, I; t)\, d\theta + \widetilde{\widetilde{H}}^1(\theta, I; t).\end{aligned}$$

Moreover, by taking our initial choice of nonresonance parameter C sufficiently small, we can ensure that the action domain D_3 contains a collection of nonresonant actions $E_\kappa(t)$, with

$$\nabla_I(\widetilde{\widetilde{H}}^0(E_\kappa(t))) = \Omega_\kappa, \tag{3E.19}$$

where

$$\widetilde{\widetilde{H}}^0(I; t) = H^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} H^1(\theta, I)\, d\theta + (2\pi)^{-n} \int_{\mathbb{T}^n} \widetilde{H}^1(\theta, I; t)\, d\theta. \tag{3E.20}$$

We now summarise the preceding discussion.

Proposition 3.12. *Suppose $H(\theta, I; t)$ is a family of real-analytic perturbations of the completely integrable nondegenerate Hamiltonian $H^0(I)$ in $\mathbb{T}^n \times D \times (-1, 1)$ that has an analytic extension to*

$$W_{s,r}(D) := \{(\theta, I) \in \mathbb{C}^n/(2\pi\mathbb{Z}) \times \mathbb{C}^n : |\mathrm{Im}(\theta)| < s, \mathrm{dist}(I, D) < r\}. \tag{3E.21}$$

Suppose further that the conditions

$$\left|\frac{\partial H^0}{\partial I}\right| \leq E, \tag{3E.22}$$

$$\left|\left(\frac{\partial^2 H^0}{\partial I^2}\right)^{-1}\right| \leq \eta, \tag{3E.23}$$

$$\left(\left|\frac{\partial H^1}{\partial I}\right| + r^{-1}\left|\frac{\partial H^1}{\partial \theta}\right|\right) \leq \epsilon \tag{3E.24}$$

are satisfied.

Then for sufficiently small $\delta > 0$, *there exists a subdomain* $\widetilde{D} \subset D$ *and a family of real-analytic symplectic maps*

$$\chi : \mathbb{T}^n \times \widetilde{D} \times (-\delta, \delta) \to \mathbb{T}^n \times D \tag{3E.25}$$

that analytically extend to a new domain of holomorphy

$$W_{s_+,r_+}(\widetilde{D}) \tag{3E.26}$$

such that

$$(H \circ \chi)(\theta, I; t) = \widetilde{H}^0(I; t) + \widetilde{H}^1(\theta, I; t), \tag{3E.27}$$

with

$$\partial_t \widetilde{H}^0(I; 0) = (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta \tag{3E.28}$$

and

$$|\widetilde{H}^1|_{s_+,r_+} = O(t^{9/4}), \tag{3E.29}$$

with constant depending only on n *and* E. *Moreover, this domain* $\widetilde{D}$ *contains a collection* $E_\kappa(t)$ *of actions such that*

$$\nabla_I(\widetilde{H}^0)(E_\kappa(t)) = \Omega_\kappa. \tag{3E.30}$$

We can also generalise this result to the Gevrey setting.

Proposition 3.13. *Suppose* $H(\theta, I; t) \in G^{\rho,\rho,1}(\mathbb{T}^n \times D \times (-1, 1)$ *is a family of Hamiltonians satisfying the assumptions of Theorem 3.10, where* $H^0(I) := H(\theta, I; 0)$ *for fixed* $\rho > 1$, *and choose* $\kappa > 0$ *small. Then for sufficiently small* $\|H(\theta, I; t) - H^0(I)\|_{L_1,L_2,L_2}$, *there exists a subdomain* $\widetilde{D} \subset D$ *and a* $G^{\rho,\rho,1}$ *family of symplectic maps*

$$\chi : \mathbb{T}^n \times \widetilde{D} \times (-1, 1) \to \mathbb{T}^n \times D \tag{3E.31}$$

such that

$$(H \circ \chi)(\theta, I; t) = \widetilde{H}^0(I; t) + \widetilde{H}^1(\theta, I; t), \tag{3E.32}$$

with

$$\partial_t \widetilde{H}^0(I; 0) = (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta \tag{3E.33}$$

and

$$\|\widetilde{H}^1\|_{CL_1,CL_2,CL_2} = O(t^{9/4}), \tag{3E.34}$$

with constant independent of κ *and with* C *dependent only on* n *and* ρ.

Moreover, the domain $\widetilde{D}$ *contains* $E_\kappa(t) = \omega^{-1}(\Omega_\kappa; t) = (\nabla_I \widetilde{H}^0)^{-1}(\Omega_\kappa; t)$.

Proof. This result is established via the approximation of Gevrey functions by real-analytic functions. First, we define

$$H^0(I) = H(\theta, I; 0), \tag{3E.35}$$

$$H^1(\theta, I; t) = H(\theta, I; t) - H(\theta, I; 0) = \int_0^t \partial_t H(\theta, I; s)\, ds \tag{3E.36}$$

and use Proposition 3.8 to boundedly extend H^0 and H^1 to the domain $\mathbb{T}^n \times \mathbb{R}^n \times (-1, 1)$, before cutting off in I to a ball $B_{\widetilde{R}}$ with $D \subset B_{\widetilde{R}-1}$. From the same methods used in the proof of Proposition 3.4, we may then construct sequences of real-analytic functions P_j^0 and P_j^1 on shrinking j-dependent complex domains U_j containing $\mathbb{T}^n \times \mathbb{R}^n \times (-1, 1)$ with a corresponding sequence $u_j \to 0$ such that

$$|P_{j+1}^k - P_j^k|_{U_{j+1}} \le C(D^0, L_1, L_2) \exp\bigl(-\tfrac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho-1)}\bigr) \|H^k\|, \tag{3E.37}$$

$$|\partial_x^\alpha (P_j^k - H^k)(\theta, I; t)| \le C(D^0, L_1, L_2) \exp\bigl(-\tfrac{3}{4}(\rho - 1)(2L_1 u_j)^{-1/(\rho-1)}\bigr) \tag{3E.38}$$

in $\mathbb{T}^n \times B_{\widetilde{R}} \times (-1, 1)$ for $|\alpha| \le 1$. These sequences P_j^k are convergent in $G^{\rho,\rho,1}(\mathbb{T}^n \times \mathbb{R}^n \times (-1, 1))$, as is shown in [Hou and Popov 2016, Proposition 2.2]. (This fact can be readily obtained by applying Cauchy estimates to (3E.37).)

Now for each $j \in \mathbb{N}$, we can carry out the first KAM step for the real-analytic Hamiltonian $P_j = P_j^0 + P_j^1$ to obtain a real-analytic symplectic map

$$\chi_j : \mathbb{T}^n \times D_1 \to \mathbb{T}^n \times D \tag{3E.39}$$

defined in shrinking holomorphy domains such that

$$((P_j^0 + P_j^1) \circ \chi_j)(\theta, I; t) = P^0(I) + t \cdot (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t P_j^1(\theta, I; 0)\, d\theta + \widetilde{P}^1(\theta, I; t), \tag{3E.40}$$

with $\|P_j^1\| = O(t^{3/2})$. Note that for an individual KAM step, the symplectic map χ_j is defined using a generating function Φ_j that is a weighted sum of finitely many Fourier components of P_j^1; see (3E.8) and [Gallavotti 1983, (3.14)]. This implies that as $P_j^0 + P_j^1 \to H^0 + H^1$ in $G^{\rho,\rho,1}(\mathbb{T}^n \times D_1 \times (-1, 1))$, the generating functions Φ_j converge to some

$$\Phi \in G^{\rho,\rho,1}(\mathbb{T}^n \times D_1 \times (-1, 1)) \tag{3E.41}$$

in the $G^{\rho,\rho,1}$ sense. From Corollary A.5, it follows that the corresponding symplectic maps χ_j converge to some

$$\chi^1 \in G^{\rho,\rho,1}(\mathbb{T}^n \times D_1 \times (-1, 1)) \tag{3E.42}$$

in the Gevrey sense.

Similarly, the symplectic maps $\tilde{\chi}_j$ that comprise a single KAM step for the Hamiltonians

$$(P_j^0 + P_j^1) \circ \chi_j \tag{3E.43}$$

can also be seen to converge to some

$$\chi^2 \in G^{\rho,\rho,1}(\mathbb{T}^n \times D_2, \mathbb{T}^n \times D_1). \tag{3E.44}$$

It follows that the family of symplectic maps $\chi_j \circ \tilde{\chi}_j$ whose existence is asserted by applying Proposition 3.12 to $P_j^0 + P_j^1$ converge to some $\chi := \chi^1 \circ \chi^2$ in the $G^{\rho,\rho,1}$-sense. Moreover, if we write

$$(P_j^0 + P_j^1) \circ \chi_j \circ \tilde{\chi}_j = \widetilde{H}_j^0(I;t) + \widetilde{H}_j^1(\theta, I; t), \tag{3E.45}$$

in the notation of Proposition 3.12, we have that $\widetilde{H}_j^k$ are convergent sequences in $G^{\rho,\rho,1}$, and so it follows that their limits $\widetilde{H}^0$, $\widetilde{H}^1$ satisfy

$$\partial_t \widetilde{H}^0(I;0) = (2\pi)^{-n} \int_{\mathbb{T}^n} \partial_t H(\theta, I; 0)\, d\theta \tag{3E.46}$$

and

$$\|\widetilde{H}^1\|_{CL_1,CL_2,CL_2} = O(t^{9/4}) \tag{3E.47}$$

as required. $\square$

Finally, we complete our computation of $\partial_t K_0(I;0)$ for a given Hamiltonian $H(\theta, I; t)$ satisfying the conditions of Theorem 3.10 by applying Proposition 3.13 to H, prior to applying Theorem 4.1 to compute the Birkhoff normal form of the transformed Hamiltonian $\widetilde{H}(\theta, I; t)$.

By applying Proposition 3.13 to $H(\theta, I; t)$ with $\|H(\theta, I; t) - H(\theta, I; 0)\|$ sufficiently small, we can then apply Theorem 3.10 to the Hamiltonian

$$\widetilde{H}(\theta, I; t) = \widetilde{H}^0(I;t) + \widetilde{H}^1(\theta, I; t), \tag{3E.48}$$

with an improved error term.

Proposition 3.14. *Suppose the assumptions of Theorem 3.9 hold for the Hamiltonian*

$$H(\theta, I; t) \in G^{\rho,\rho,1}(\mathbb{T}^n \times D \times (-1,1)). \tag{3E.49}$$

Then there exist $N(n,\rho,\tau) > 0$ and $\epsilon > 0$ independent of L_1, L_2, D such that for any

$$H \in G^{\rho,\rho,1}_{L_1,L_2,L_2}(\mathbb{T}^n \times D \times (-1,1)),$$

with

$$\kappa^{-2}\|H(\theta, I; t) - H(\theta, I; 0)\|_{L_1,L_2,L_2} = \epsilon_H \le \epsilon L_1^{-N-2(\tau+2)}, \tag{3E.50}$$

there is a subdomain $\widetilde{D} \subset D$ containing $E_\kappa(0)$ and a family of $G^{\rho',\rho'}$ maps $\omega : \widetilde{D} \times \left(\frac{1}{2}, 1, 2\right) \to \Omega$ and a family of maps $\chi \in G^{\rho,\rho',\rho'}(\mathbb{T}^n \times \widetilde{D} \times \left(\frac{1}{2}, 1, 2\right), \mathbb{T}^n \times \widetilde{D})$ that are diffeomorphisms and exact symplectomorphisms respectively for each fixed $t \in \left(\frac{1}{2}, 1, 2\right)$. Moreover, we can choose the maps ω and χ such that family of transformed Hamiltonians

$$\widetilde{H}(\theta, I; t) := (H \circ \chi)(\theta, I; t) \tag{3E.51}$$

is of Gevrey class $G^{\rho,\rho',\rho'}\left(\mathbb{T}^n \times \widetilde{D} \times \left(\frac{1}{2}, 1, 2\right)\right)$ and can be decomposed as

$$K(I;t) + R(\theta, I; t) := \widetilde{H}(0, I; t) + (\widetilde{H}(\theta, I; t) - \widetilde{H}(0, I; t)) \tag{3E.52}$$

such that:

(i) *$\mathbb{T}^n \times \{I\}$ is an invariant Lagrangian torus of $\widetilde{H}(\cdot\,,\cdot\,; t)$ for each $I \in E_\kappa(t) = \omega^{-1}(\widetilde{\Omega}_\kappa)$ and each $t \in \left(\frac{1}{2}, 1, 2\right)$.*

(ii) $\partial_I^\beta(\nabla K(I;t)-\omega(I;t)) = \partial_I^\beta R(\theta,I;t) = 0$ *for all* $(\theta,I;t)\in\mathbb{T}^n\times E_\kappa(t)\times\left(\frac{1}{2},1,2\right)$, $\beta\in\mathbb{N}^n$.

(iii) *There exist* $A, C>0$ *independent of* κ, L_1, L_2, *and* $D\subset D^0$ *such that we have the estimates*

$$|\partial_\theta^\alpha\partial_I^\beta\partial_t^\delta\phi(\theta,I;t)| + |\partial_I^\beta\partial_t^\delta(\omega(I;t)-\nabla\widetilde{H}^0(I;t))| + |\partial_\theta^\alpha\partial_I^\beta\partial_t^\delta(\widetilde{H}(\theta,I;t)-\widetilde{H}^0(I;t))| \\ \le AC^{|\alpha|+|\beta|+|\delta|}L_1^{|\alpha|}(L_1^{\tau+1}/\kappa)^{|\beta|}\alpha!^\rho\,\beta!^{\rho'}\,\delta!^{\rho'}\,L_1^{N/2}|t|^{9/8} \tag{3E.53}$$

uniformly in $\mathbb{T}^n\times\widetilde{D}\times\left(\frac{1}{2},1,2\right)$ *for all* α,β, *where* $\phi\in G^{\rho,\rho',\rho'}\left(\mathbb{T}^n\times\widetilde{D}\times\left(\frac{1}{2},1,2\right)\right)$ *is such that* $\langle\theta,I\rangle+\phi(\theta,I;t)$ *generates the symplectomorphisms* χ *in the sense of* (3E.8) *and* $\widetilde{H}^0$, $\widetilde{H}^1$ *are as in Proposition 3.13.*

(iv) $$\partial_t K(I;t) = (2\pi)^{-n}\int_{\mathbb{T}^n}\partial_t H(\theta,I;0)+o(1) \tag{3E.54}$$

uniformly in $\mathbb{T}^n\times\widetilde{D}\times\left(\frac{1}{2},1,2\right)$.

Proof. The only new claim in this proposition is (3E.54), which follows from (3E.53) and the expression (3E.33) for $\widetilde{H}^0$. Note that the exponent $\frac{9}{8}$ in (3E.53) comes from (3E.34) and the square root in (3D.7). □

4. Quantum Birkhoff normal form

Through the work in Section 3, we have now established that the Birkhoff normal form construction in [Popov 2004b] preserves smoothness in the t-parameter when applied to the Hamiltonian $P_0(x,\xi;t)$ that is the principal symbol of the operator introduced in (1B.1). This regularity in t propagates through the quantum Birkhoff normal form construction in [Popov 2004a], which we discuss in this section. The upshot of this regularity in t is that the quasimodes constructed in [Popov 2004a, Section 2.4] can be chosen to have associated quasieigenvalues varying smoothly in the parameter t. We discuss these quasimodes in Section 4C.

4A. ***Quantum Birkhoff normal form.*** In [Popov 2004a], a quantum Birkhoff normal form is constructed for semiclassical pseudodifferential operators of the form (1B.1) after first obtaining a classical Birkhoff normal form for the principal symbol of regularity $G^{\rho,\rho'}$ as in Theorem 3.10. This normal form uses the Gevrey symbol classes introduced in Section B and is stated in Theorem 4.1. We remark that the proof is presented in [Popov 2004a] for differential operators, but can be carried out without change if the $\mathcal{P}_h$ is a pseudodifferential operator.

We denote by χ_1 the symplectomorphism that transforms the completely integrable Hamiltonian $P(x,\xi;0)$ into action-angle coordinates $H = P\circ(\chi_1)$ and we denote by $\chi_0(t)$ the symplectomorphism that transforms the perturbed Hamiltonian $H(\theta,I;t)$ into Birkhoff normal form, as constructed in Theorem 3.10. For the purpose of stating the quantum Birkhoff normal form for $\mathcal{P}_h(t)$, the Maslov class of the KAM tori $\Lambda_\omega:\omega\in\Omega_\kappa\}$ (as defined in Section 3.4 of [Duistermaat 1996]) can be identified with elements of $\vartheta\in H^1(\mathbb{T}^n;\mathbb{Z})$ via the family of symplectomorphisms $\chi_0(t)\circ\chi_1:\mathbb{T}^n\times D\to T^*M$. Following [Popov 2000b; Colin de Verdière 1977], we can then associate a smooth line bundle $\mathbb{L}$ over $\mathbb{T}^n$ with the class ϑ such that smooth sections $f\in C^\infty(\mathbb{T}^n,\mathbb{L})$ can be canonically identified with smooth functions

$\tilde{f} \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{C})$ satisfying the quasiperiodicity condition

$$\tilde{f}(x + 2\pi p) = \exp\left(\frac{i\pi}{2}\langle\vartheta, p\rangle\right)\tilde{f}(x) \tag{4A.1}$$

for all $p \in \mathbb{Z}^n$.

The quantum Birkhoff normal form in [Popov 2004a] is far sharper than is necessary for the purposes of this paper, with remainders of order $O(e^{-ch^{-1/\nu}})$. We require only the following truncated version, with error terms of order $O(h^{\gamma+1})$ for some fixed $\gamma > 0$.

Theorem 4.1. *Suppose $\mathcal{P}_h(t)$ is as in (1B.1). Then for each fixed t there exists a uniformly bounded family of semiclassical Fourier integral operators*

$$U_h(t) : L^2(\mathbb{T}^n; \mathbb{L}) \to L^2(M), \quad 0 < h < h_0, \tag{4A.2}$$

that are associated with the canonical relation graph of the Birkhoff normal form transformation $\chi(t)$ such that we have

(i) $U_h(t)^* U_h(t) - \mathrm{Id}$ *is a pseudodifferential operator with symbol in the Gevrey class $S_\ell(\mathbb{T}^n \times D)$ which restricts to an element of $h^{\gamma+1} S_\ell(\mathbb{T}^n \times Y)$ for some subdomain Y of D that contains $E_\kappa(t)$,*

(ii) $\mathcal{P}_h(t) \circ U_h(t) - U_h(t) \circ \mathcal{P}_h^0(t) = \mathcal{R}_h(t) \in h^{\gamma+1} S_\ell$, *where the operator $\mathcal{P}_h^0(t)$ has symbol*

$$p^0(\theta, I; t, h) = K^0(I; t, h) + R^0(\theta, I; t, h) = \sum_{j \le \gamma} K_j(I; t) h^j + \sum_{j \le \gamma} R_j(\theta, I; t) h^j, \tag{4A.3}$$

with both K^0 and R^0 in the symbol class $S_\ell(\mathbb{T}^n \times D)$ from Definition B.5 where $\eta > 0$ is a constant, $K_0(I; t)$, $R_0(\theta, I; t)$ are the components of the Birkhoff normal form of the Hamiltonian $P_0 \circ \chi_1$ as constructed in Theorem 3.10, and

$$\partial_I^\alpha R_j(\theta, I; t) = 0 \tag{4A.4}$$

for $(\theta, I; t) \in \mathbb{T}^n \times E_\kappa(t) \times (-1, 1)$. Moreover, the symbols K_j, R_j in (4A.3) are smooth in the parameter t.

Our statement of Theorem 4.1 differs from [Popov 2004a, Theorem 2.1] only in the presence of the parameter t, the smoothness of the symbols K_j, R_j in t, and the truncation to fixed finite order $O(h^{\gamma+1})$. We sketch the details of the proof of Theorem 4.1 in this section, following the argument of [Popov 2004a].

The construction of $U_h(t)$ can be broken into multiple steps. We begin by constructing a family of semiclassical Fourier integral operators $T_h(t)$ that conjugate $P_h(t)$ to a family of semiclassical pseudodifferential operators $P_h^1(t) : \mathcal{C}^\infty(\mathbb{T}^n; \mathbb{L})$ with principal symbol equal to $K_0(I; t) + R_0(\theta, I; t)$, the Birkhoff normal form of H, and with vanishing subprincipal symbol. The conjugating semiclassical Fourier integral operators arise by quantising the G^ρ symplectomorphisms

$$\chi_1 : \mathbb{T}^n \times D \to T^*M, \tag{4A.5}$$

$$\chi_0 : \mathbb{T}^n \times D \to \mathbb{T}^n \times D \tag{4A.6}$$

that transform the unperturbed Hamiltonian $P(x, \xi; 0)$ to action-angle variables and transform the perturbed Hamiltonian to Birkhoff normal form respectively, and composing these two operators. Full details for this construction can be found in [Popov 2000b, Section 2].

From the regularity of the symplectomorphisms, it follows that there exists a semiclassical expansion for $P_h^1(t)$ with symbols smooth in t.

The symbol of the operator $P_h^1(t)$ satisfies the property (4A.3) to $O(h^2)$, and to improve this, we replace the conjugating Fourier integral operator T_h with $T_h A_h$ for a suitable elliptic pseudodifferential operator A_h whose symbol is determined iteratively on the family of Cantor-like sets

$$\{(\theta, I; t) \in \mathbb{T}^n \times \mathbb{R}^n \times (-1, 1) : I \in E_\kappa(t)\}$$

by solving equations of the form

$$\langle \nabla K_0, \partial_\theta \rangle f(\theta, I; t) = g(\theta, I; t), \tag{4A.7}$$

referred to in the literature as homological equations. In this manner the "flatness condition" of (4A.4) is obtained for $j > 0$, where the $j = 0$ statement is established by Theorem 3.10. We outline this procedure in Section 4B.

The key fact is that the homological equation can be solved smoothly in the parameter t, which is the content of Theorem 4.3. One can then apply Theorem 4.3 as in [Popov 2004a, Section 2.3] to complete the construction of the quantum Birkhoff normal form, with the additional consequence of smoothness of symbols K_j, R_j.

4B. ***Construction of the quantum Birkhoff normal form.*** After conjugating $P_h(t)$ by semiclassical Fourier integral operators as described in the previous section, we obtain a family of self-adjoint semiclassical operators $P_h^1(t)$ with symbol $\tilde{p} \in S_{\tilde{\ell}}(\mathbb{T}^n \times D)$ satisfying the flatness condition (4A.4) to order h^2, where $\tilde{\ell} = (\rho, \rho', \rho + \rho' - 1)$. That is to say, the formal summation of $\tilde{p}$,

$$\sum_{j=0}^{\infty} \tilde{p}_j(\theta, I; t) h^j, \tag{4B.1}$$

satisfies

$$\tilde{p}_0(\theta, I; t) = K_0(I; t) + R_0(\theta, I; t), \tag{4B.2}$$

$$\tilde{p}_1(\theta, I; t) = 0. \tag{4B.3}$$

The next step of the proof of Theorem 4.1 is the improvement of the order of the flatness condition by composition with a suitable elliptic semiclassical pseudodifferential operator

$$A_h(t) = \mathrm{Id} + O(h)$$

with symbol

$$a(\theta, I; t) = \sum_{j=1}^{\infty} a_j(\theta, I; t) h^j. \tag{4B.4}$$

To motivate the method, we suppose that a quantum Birkhoff normal form P_h^0 exists in the sense of Theorem 4.1. Our current operator $\widetilde{P}_h$ is equal to P_h^0 up to order h^2 by construction. Hence, we have

$$\begin{aligned} T_h(t) A_h(t) \widetilde{P}_h(t) &= T_h(t) \widetilde{P}_h(t) A_h(t) + T_h(t)[A_h(t), \widetilde{P}_h(t)] \\ &= P_h^1(t) T_h(t) A_h(t) + h^2 T(t) B(t) A(t) + T_h(t)[A_h(t), \widetilde{P}_h(t)] \end{aligned} \tag{4B.5}$$

for some semiclassical pseudodifferential operator $B_h(t)$ in the symbol class $S_{\tilde{\ell}}(\mathbb{T}^n \times D)$. From composition formulae, the symbol of the commutator is equal to

$$-(\partial_\theta^\alpha a_1 \partial_I^\alpha \tilde{p}_0)h^2 = -\mathcal{L}_{\omega I;t} a_1, \tag{4B.6}$$

where $\mathcal{L}_\omega = \langle \omega, \partial_\theta \rangle a_1(\theta, I; t)$. Thus to improve the order of the flatness condition, it suffices to choose a_1 solving the homological equation

$$\mathcal{L}_{\omega(I;t)} a_1 = b_0, \tag{4B.7}$$

where b_0 denotes the principal symbol of $B_h(t)$. Indeed, if (4B.7) is solvable, then we have

$$T_h(t) A_h(t) P_h(t) = P_h^0(t) T_h(t) A_h(t) + O(h^3). \tag{4B.8}$$

Extending this idea, it was shown in [Popov 2000b] that we can choose higher-order terms of the symbol a in an iterative fashion by the solution of such a homological equation for each power of h that we gain. The consequence is the following result.

Proposition 4.2. *There exist $a, K^0, r \in S_\ell(\mathbb{T}^n \times D)$, where $\ell = (\rho, \mu, \nu)$, such that*

$$a(\theta, I; t, h) \sim \sum_{j=0}^\infty a_j(\theta, I; t) h^j, \tag{4B.9}$$

$$K^0(I; t, h) \sim \sum_{j=0}^\infty K_j(I; t) h^j, \tag{4B.10}$$

$$r(\theta, I; t, h) \sim \sum_{j=0}^\infty r_j(\theta, I; t) h^j, \tag{4B.11}$$

where $a_0 = 1,\ r_0 = R_0,\ K_1 = 0$, and

$$\tilde{p} \circ a - a \circ K^0 \sim r, \tag{4B.12}$$

where each $r_j(\theta, I; t)$ is flat in I on $\mathbb{T}^n \times E_\kappa(t)$.

The symbol K^0 in the statement of theorem corresponds to the sought symbol K^0 in Theorem 4.1, while the symbol R^0 is then constructed by solving $a \circ R^0 = r$, which is possible by ellipticity.

The completion of the proof of Theorem 4.1 after establishing Proposition 4.2 is contained in [Popov 2000b, Section 3]. For our additional requirement of smoothness in t in Theorem 4.1, it thus suffices to verify that the homological equation can be solved smoothly in the parameter t. In particular, we require the following.

Theorem 4.3. *Suppose $f(\cdot, \cdot; t) \in G^{\rho,\mu}(\mathbb{T}^n \times D)$ satisfies the estimate*

$$|\partial_\theta^\alpha \partial_I^\beta f(\theta, I; t)| \le d_0 C^{|\alpha|+\mu|\beta|} \Gamma(\rho|\alpha| + \mu|\beta| + q) \tag{4B.13}$$

uniformly in the smooth parameter $t \in (-1, 1)$ for some $q > 0$ and some $C \ge 1$ and that for each $I \in D$, we have

$$\int_{\mathbb{T}^n} f(\theta, I; t)\, d\theta = 0. \tag{4B.14}$$

Then for any smooth family $\omega(\,\cdot\,;t) \in G^{\rho'}_{L_0}(D,\Omega)$ there is a solution $u(\,\cdot\,,\cdot\,;t) \in G^{\rho,\mu}(\mathbb{T}^n \times D)$ to the equation

$$\mathcal{L}_\omega u(\theta, I; t) = f(\theta, I; t), \quad (\theta, I) \in \mathbb{T}^n \times E_\kappa(t), \tag{4B.15}$$

$$u(0, I; t) = 0, \qquad I \in D, \tag{4B.16}$$

where $\mathcal{L}_\omega = \langle \omega(I;t), \partial/\partial\theta\rangle$. Moreover, u is smooth in the parameter t and satisfies the estimate

$$|\partial_\theta^\alpha \partial_I^\beta u(\theta, I; t)| \leq A d_0 C^{n+\tau+|\alpha|+\mu|\beta|+1}\Gamma(\rho|\alpha| + \mu|\beta| + \rho(n+\tau+1)+q), \tag{4B.17}$$

where A depends only on n, ρ, τ and μ.

This theorem statement differs from [Popov 2004a, Proposition 2.3] only in the presence of the smooth parameter t, and indeed an identical proof based on taking the Fourier expansion

$$u(\theta, I; t) = \sum_{k \in \mathbb{Z}^n} e^{i\langle k,\theta\rangle} u_k(I;t) \tag{4B.18}$$

and solving for u_k can be pursued. The rapid decay of Fourier coefficients established in [Popov 2004a] implies that the limit $u(\theta, I; t)$ is smooth in t as required. The proof is then identical to that in [Popov 2004a], with the uniformity in (4B.17) following from the uniformity in (4B.13).

4C. *Quasimode construction.* We now briefly outline how the construction of Gevrey class quasimodes for $\mathcal{P}_h(t)$ follow from the quantum Birkhoff normal form Theorem 4.1. These quasimodes microlocalise onto a family of nonresonant tori and moreover have quasieigenvalues that are smooth in the parameter $t \in (-1, 1)$.

Definition 4.4. An $O(h^{\gamma+1})$ family of G^ρ quasimodes $\mathcal{Q}(t)$ for $\mathcal{P}_h(t)$ is a family

$$\{(v_m(x;t,h), \mu_m(t,h)) : m \in \mathcal{M}_h(t)\} \subset \mathcal{C}^\infty(M \times \mathcal{D}_h(m)) \times \mathcal{C}^\infty(\mathcal{D}_h(m)) \tag{4C.1}$$

parametrised by $h \in (0, h_0]$, where

- $\mathcal{M}_h(t) \subset \mathbb{Z}^n$ is an h-dependent finite index set,
- $\mathcal{D}_h(m) = \{t \in (-1,1) : m \in \mathcal{M}_h(t)\}$,
- each $v_m(\,\cdot\,;t,h)$ is uniformly of class G^ρ,
- $\|P_h(t)v_m(\,\cdot\,;t,h) - \mu_m(t;h)v_m(\,\cdot\,;t,h)\|_{L^2} = O(h^{\gamma+1})$ for all $m \in \mathcal{M}_h(t)$,
- $|\langle v_m(\,\cdot\,;t,h), v_l(\,\cdot\,;t,h)\rangle - \delta_{ml}| = O(h^{\gamma+1})$ for all $m, l \in \mathcal{M}_h(t)$.

Theorem 4.5. *Suppose now that $t \in (-1,1)$ is fixed and $S \subset E_\kappa(t)$ is a closed collection of nonresonant actions. For an arbitrary constant $L > 1$, we define the index set*

$$\mathcal{M}_h := \{m \in \mathbb{Z}^n : \operatorname{dist}(S, h(m + \vartheta/4)) < Lh\}, \tag{4C.2}$$

where $\vartheta \in \mathbb{Z}^n$ is the Maslov class of any Lagrangian tori $\{\chi(\mathbb{T}^n \times \{I\})\}$ with $I \in S$. Note that this class is independent of the choice of torus by the local constancy of the Maslov class.

Then

$$\{(v_m(x;t,h),\mu_m(t;h)) : m \in \mathcal{M}_h(t)\} := (U_h(t)e_m, K^0(h(m+\vartheta/4);t,h) \tag{4C.3}$$

defines a G^ρ family of quasimodes for $\mathcal{P}_h(t)$ that has Gevrey microsupport on the family of tori

$$\Lambda_S = \bigcup_{I\in S} \Lambda_{\omega(I;t)} = \bigcup_{I\in S} \chi(\mathbb{T}^n \times \{I\}) \subset T^*M, \tag{4C.4}$$

where $\{e_m\}_{m\in\mathbb{Z}^n}$ is the orthonormal basis of $L^2(\mathbb{T}^n;\mathbb{L})$ associated to the quasiperiodic functions

$$\tilde{e}_m(x) := \exp(i\langle m+\vartheta/4, x\rangle). \tag{4C.5}$$

Proof. From the definition of the functions e_m, it follows that

$$\begin{aligned} P_h^0(t)(e_m)(\theta) &= \sigma(P_h^0(t))(\theta, h(m+\vartheta/4))e_m(\theta) \\ &= (K^0(h(m+\vartheta/4);t,h) + R^0(\theta, h(m+\vartheta/4);t,h))e_m(\theta) \\ &= (\lambda_m(t;h) + R^0(\theta, h(m+\vartheta/4))e_m(\theta). \end{aligned} \tag{4C.6}$$

From the definition (4C.2) of the index set $\mathcal{M}_h(t)$ and from $A.2$, it thus follows that

$$\mathcal{P}_h(t)(U_h(t)e_m) = U_h(t)P_h^0(t)e_m = O(h^{\gamma+1}) \tag{4C.7}$$

upon an application of Theorem 4.1. The almost-orthogonality of the $U_h(t)e_m$ then follows from the fact that $U_h(t)$ is almost unitary by Theorem 4.1, and that the e_m are exactly orthogonal by construction. □

These quasimodes are as numerous as we could hope for; indeed the index set $\mathcal{M}_h(t)$ satisfies the local Weyl asymptotic

$$\lim_{h\to 0}(2\pi h)^n \,\#\,\mathcal{M}_h = m(\mathbb{T}^n \times S) = \mu(\Lambda_S), \tag{4C.8}$$

where m denotes the $(2n)$-dimensional Lebesgue measure and μ denotes the symplectic measure $d\xi\, dx$. To see this, we can denote by U the union of n-cubes centred at the lattice points in $\mathcal{M}_h$ with side length h. The containment

$$S \subset U \subset \{I : \operatorname{dist}(I,S) < \tilde{L}h\} \tag{4C.9}$$

for a constant $\tilde{L}$ then yields the claim by monotone convergence of measures, noting that since S is closed we have

$$S = \bar{S} = \bigcap_{h>0} \{I : \operatorname{dist}(I,S) < \tilde{L}h\}. \tag{4C.10}$$

In the special case of $S = \{I\}$, we have a family of G^ρ quasimodes with microsupport on an individual torus $\chi(\mathbb{T}^n \times \{I\})$.

Appendix A: Anisotropic Gevrey classes

In this appendix, we define the Gevrey function spaces used throughout the paper and collect several of their properties from the appendix of [Popov 2004b].

Definition A.1. For $\rho \geq 1$ and $X \subset \mathbb{R}^n$ open, the Gevrey class of order ρ is given by

$$G_L^\rho(X) := \big\{ f \in \mathcal{C}^\infty(X) : \sup_\alpha \sup_{x \in X} |\partial_x^\alpha f(x)| L^{-|\alpha|} \alpha!^{-\rho} < \infty \big\}. \tag{A.1}$$

If $f \in G_L^\rho(X)$, the supremum in (A.1) is denoted by $\|f\|_L$. We will frequently suppress the L in our notation. Equipped with this norm, $G_L^\rho(X)$ is a Banach space. Gevrey regularity is generally weaker the real-analyticity (they coincide when $\rho = 1$ as can be seen by using the Cauchy–Hadamard theorem to characterise analytic functions by the growth of their Taylor coefficients) and importantly, there exist bump functions in the Gevrey class for $\rho > 1$.

An important property of the Gevrey class that follows from Taylor's theorem is that if a Gevrey function has vanishing derivatives, then locally it is superexponentially small.

Proposition A.2. *Suppose $f \in G^\rho(X)$, and $\rho > 1$. Then there exist positive constants c, C, η and r_0 only dependent on the Gevrey constant L, the norm $\|f\|_L$, and the set X such that*

$$f(x_0 + r) = \sum_{|\alpha| \leq \eta |r|^{1/(1-\rho)}} f_\alpha(x_0) r^\alpha + R(x_0, r), \tag{A.2}$$

where $f_\alpha = (\partial^\alpha f)/\alpha!$ and

$$|\partial_x^\beta R(x_0, r)| \leq C^{1+|\beta|} \beta!^\rho \, e^{-c|r|^{-1/(\rho-1)}} \quad \textit{for all } 0 < |r| \leq \min(r_0, d(x_0, \mathbb{R}^n \setminus X)). \tag{A.3}$$

We also need to consider anisotropic Gevrey classes, which are classes of Gevrey functions with differing regularity in individual variables.

Definition A.3. Suppose X and Y are open subsets of Euclidean spaces. Suppose that $\rho_1, \rho_2 \geq 1$ and $L_1, L_2 > 0$. Then

$$G_{L_1,L_2}^{\rho_1,\rho_2}(X \times Y) = \big\{ f \in \mathcal{C}^\infty(X \times Y) : \sup_{(x,y) \in X \times Y} |\partial_x^\alpha \partial_y^\beta f| L_1^{-|\alpha|} L_2^{-|\beta|} \alpha!^{-\rho_1} \beta!^{-\rho_2} < \infty \big\}. \tag{A.4}$$

If $f \in G_{L_1,L_2}^{\rho_1,\rho_2}$, then we denote the supremum in (A.4) by $\|f\|_{L_1,L_2}$. Equipped with this norm, $G_{L_1,L_2}^{\rho_1,\rho_2}$ is a Banach space. This definition extends in the natural way to $k \geq 3$ variables. Furthermore, some of these variables might lie in complex domains.

In anisotropic Gevrey classes, one has the following implicit function theorem due to Komatsu.

Proposition A.4. *Suppose that $F \in G_{L_1,L_2}^{\rho,\rho'}(X \times \Omega^0, \mathbb{R}^n)$, where $X \subset \mathbb{R}^n$, $\Omega^0 \subset \mathbb{R}^m$ and*

$$L_1 \|F(x, \omega) - x\|_{L_1,L_2} \leq \tfrac{1}{2}.$$

Then there exists a local solution $x = g(y, \omega)$ to the implicit equation

$$F(x, \omega) = y \tag{A.5}$$

defined in a domain $Y \times \Omega$. Moreover, there exist constants A, C dependent only on ρ, ρ', n, m such that $g \in G_{CL_1,CL_2}^{\rho,\rho'}(Y \times \Omega, X)$, with $\|g\|_{CL_1,CL_2} \leq A \|F\|_{L_1,L_2}$.

A consequence of this theorem is established in [Popov 2004b].

Corollary A.5. *Suppose $F \in G^{\rho,\rho'}_{L_1,L_2}(\mathbb{T}^n \times \Omega, \mathbb{T}^n)$, where $\Omega^0 \subset \mathbb{R}^m$ and $L_1\|F(\theta,\omega)-\theta\|_{L_1,L_2} \le \frac{1}{2}$. Then there exists a local solution $x = g(y,\omega)$ to the implicit equation*

$$F(x,\omega) = y \tag{A.6}$$

defined on $\mathbb{T}^n \times \Omega$. Moreover, there exist positive constants A, C dependent only on ρ, ρ', n, m such that $g \in G^{\rho,\rho'}_{CL_1,CL_2}(\mathbb{T}^n \times \Omega)$ with $\|g\|_{CL_1,CL_2} \le A\|F\|_{L_1,L_2}$.

Finally, we have two results on the composition of functions of Gevrey regularity, which can also be found in [Popov 2004b].

Proposition A.6. *Let $X \subset \mathbb{R}^n$, $Y \subset \mathbb{R}^m$, and $\Omega \subset \mathbb{R}^k$ be open sets. Suppose $g \in G^{\rho'}_{L_1}(\Omega, Y)$, with $\|g\|_{L_1} = A_1$, and $f \in G^{\rho,\rho'}_{B,L_2}(X \times Y)$, with $\|f\|_{B,L_2} = A_2$. Then the composition $F(x,\omega) := f(x, g(\omega))$ is in $G^{\rho,\rho'}_{B,L}(X \times \Omega)$, where*

$$L = 2^{l+\rho'} l^{\rho'} L_1 \max(1, A_1 L_2),$$

with $l = \max(k,m,n)$. Moreover we have the Gevrey norm estimate

$$\|F\|_{B,L} \le A_2.$$

Proposition A.7. *Let $X \subset \mathbb{R}^n$, $Y \in \mathbb{R}^m$, and $\Omega \subset \mathbb{R}^k$ be open sets. Suppose $g \in G^{\rho,\rho'}_{B_1,L_1}(X \times \Omega, Y)$ with $\|g\|_{B_1,L_1} = A_1$ and $f \in G^{\rho,\rho'}_{B_2,L_2}(Y \times \Omega)$. Then the composition*

$$F(x,\omega) := f(g(x,\omega),\omega)$$

is in $G^{\rho,\rho'}_{B,L}(X \times \Omega)$, where

$$B = 4^l (4l)^\rho B_1 \max(1 + A_1 B_2),$$
$$L = L_2 + 4^l (4l)^\rho L_1 \max(1, A_1 B_2),$$

with $l = \max(k,m,n)$. Moreover we have the Gevrey norm estimate

$$\|F\|_{B,L} \le A_2.$$

Appendix B: Gevrey class symbols

In this appendix, we introduce the class of Gevrey symbols used throughout this paper. We suppose D is a bounded domain in $\mathbb{R}^n$, and take $X = \mathbb{T}^n$ or a bounded domain in $\mathbb{R}^m$. We fix the parameters $\sigma, \mu > 1$ and $\varrho \ge \sigma + \mu - 1$, and denote the triple (σ, μ, ϱ) by ℓ.

Definition B.1. A formal Gevrey symbol on $X \times D$ is a formal sum

$$\sum_{j=0}^{\infty} p_j(\theta, I) h^j, \tag{B.1}$$

where the $p_j \in \mathcal{C}_0^\infty(X \times D)$ are all supported in a fixed compact set and there exists a $C > 0$ such that

$$\sup_{X \times D} |\partial_\theta^\beta \partial_I^\alpha p_j(\theta, I)| \le C^{j+|\alpha|+|\beta|+1} \beta!^\sigma \, \alpha!^\mu \, j!^\varrho. \tag{B.2}$$

Definition B.2. A realisation of the formal symbol (B.1) is a function $p(\theta, I; h) \in \mathcal{C}_0^\infty(X \times D)$ for $0 < h \leq h_0$, with

$$\sup_{X\times D\times(0,h_0]}\left|\partial_\theta^\beta \partial_I^\alpha\left(p(\theta, I; h) - \sum_{j=0}^{N} p_j(\theta, I)h^j\right)\right| \leq h^{N+1}C_1^{N+|\alpha|+|\beta|+2}\beta!^\sigma\,\alpha!^\mu\,(N+1)!^\varrho. \tag{B.3}$$

Lemma B.3. *Given a formal symbol* (B.1), *one choice of realisation is*

$$p(\theta, I; h) := \sum_{j\leq\epsilon h^{-1/\varrho}} p_j(\theta, I)h^j, \tag{B.4}$$

where ϵ *depends only on* n *and* C_1.

Definition B.4. We define the residual class of symbols $S_\ell^{-\infty}$ as the collection of realisations of the zero formal symbol.

Definition B.5. We write $f \sim g$ if $f - g \in S_\ell^{-\infty}$. It then follows that any two realisations of the same formal symbol are $\sim$-equivalent. We denote the set of equivalence classes by $S_\ell(X \times D)$.

We now discuss the class of pseudodifferential operators corresponding to these symbols.

Definition B.6. To each symbol $p \in S_\ell(X \times D)$, we associate a semiclassical pseudodifferential operator defined by

$$(2\pi h)^{-n}\int_{X\times\mathbb{R}^n} e^{i(x-y)\cdot\xi/h}p(x, \xi; h)u(y)\,d\xi\,dy \tag{B.5}$$

for $u \in \mathcal{C}_0^\infty(X)$.

The above construction is well-defined modulo $\exp(-ch^{-1/\varrho})$, as for any $p \in S_\ell^{-\infty}(X \times D)$ we have

$$\|P_h u\| = O_{L^2}(\exp(-ch^{-1/\varrho})) \tag{B.6}$$

for some constant $c > 0$.

Remark B.7. The exponential decay of residual symbols is a key gain that comes from working in a Gevrey symbol class.

The operations of symbol composition and conjugation then correspond to composing operators and taking adjoints respectively. Moreover, if $p \in S_{(\sigma,\sigma,2\sigma-1)}$, then G^σ-smooth changes of variable preserve the symbol class of p. This coordinate invariance allows us to extend the Gevrey pseudodifferential calculus to compact Gevrey manifolds.

Appendix C: Estimates for analytic functions

In this appendix we prove several elementary but important estimates for analytic functions.

Proposition C.1. *Suppose* $\widetilde{\Omega}_j \subset \mathbb{C}$ *are open sets and* $\Omega_j \subset \widetilde{\Omega}_j$ *are such that* $\operatorname{dist}(\Omega_j, \mathbb{C}\setminus\widetilde{\Omega}_j) > r_j$. *Define*

$$\Omega = \prod_{j=1}^{n}\Omega_j \quad \textit{and} \quad \widetilde{\Omega} = \prod_{j=1}^{n}\widetilde{\Omega}_j. \tag{C.1}$$

If the analytic function $f : \widetilde{\Omega} \to \mathbb{C}$ *satisfies*

$$\|f\|_\Omega = A < \infty \tag{C.2}$$

then we have

$$\|\partial_z^\alpha f\|_\Omega \le A r^{-\alpha} \alpha! \tag{C.3}$$

for each multi-index α.

Proof. For $z \in \Omega$, the Cauchy integral formula implies

$$f(z) = \frac{1}{(2\pi i)^n} \oint_{\partial B(z_1,r_1)} \oint_{\partial B(z_2,r_2)} \cdots \oint_{\partial B(z_n,r_n)} \frac{f(w)}{w-z}\, dw_1\, dw_2 \cdots dw_n, \tag{C.4}$$

which yields

$$\partial_z^\alpha f(z) = \frac{\alpha!}{(2\pi i)^n} \oint_{\partial B(z_1,r_1)} \oint_{\partial B(z_2,r_2)} \cdots \oint_{\partial B(z_n,r_n)} \frac{f(w)}{(w-z)^{\alpha+1}}\, dw_1\, dw_2 \cdots dw_n \tag{C.5}$$

upon repeated differentiation, where 1 denotes the multi-index $(1, 1, \dots, 1)$. Hence

$$\|\partial_z^\alpha f\|_\Omega \le A r^{-\alpha} \alpha! \tag{C.6}$$

as required. □

We also have an implicit function theorem for real-analytic functions. Defining

$$O_h = \{\omega \in \mathbb{C}^n : \mathrm{dist}(\omega, \Omega) < h\}, \tag{C.7}$$

where distances in $\mathbb{C}^n$ are taken with the sup-norm, we have the following.

Proposition C.2. *Suppose* $f : O_h \times (-1, 1) \to \mathbb{C}^n$ *is real-analytic, and we have the estimate*

$$|f|_h < \infty. \tag{C.8}$$

Then, for any $0 < v < \frac{1}{6}$ *such that*

$$|f - \mathrm{id}|_h \le vh, \tag{C.9}$$

the function has a real-analytic inverse $g : O_{(1/2-3v)h} \times (-1, 1) \to O_{(1-4v)h}$ *that satisfies the estimate*

$$\max(|g - \mathrm{id}|_{(1/2-3v)h}, 3vh|D\phi - \mathrm{Id}|_{(1/2-3v)h}) \le |f - \mathrm{id}|_h \tag{C.10}$$

uniformly in $t \in (-1, 1)$. *The norm* $|\cdot|_h$ *denotes the sup-norm over* O_h *and the matrix norm in* (C.10) *is the norm induced by equipping* $\mathbb{C}^n$ *with the sup-norm.*

Proposition C.2 can be proven in the same way as in Lemma 3.4 of [Popov 2004b]. The only difference is that we need to work on domains of the form $O_{\lambda h} \times B_1^{\mathbb{C}}$, and invert maps of the form

$$\tilde{f}(\omega, t) := (f(\omega, t), t) \tag{C.11}$$

for given f satisfying the assumptions of the proposition uniformly in t.

Appendix D: Whitney extension theorem

In this appendix, we prove a version of the Whitney extension theorem for anisotropic Gevrey classes. The proof is adapted from [Bruna 1980] in the non-anisotropic case.

Definition D.1.

$$\mathcal{C}^\infty_{M,\widetilde{M}}(X\times Y)=\left\{f\in\mathcal{C}^\infty(X\times Y,\mathbb{R}):\sup_{(x,y)\in X\times Y}\sup_{\alpha,\beta}\left(\frac{|(\partial_x^\alpha\partial_y^\beta f)(x,y)|}{L_1^{|\alpha|}L_2^{|\beta|}M_{|\alpha|}\widetilde{M}_{|\beta|}}\right)<\infty\text{ for some }L_j>0\right\},\tag{D.1}$$

where X, Y are open sets in Euclidean spaces of possibly differing dimension, α, β are multi-indices of the appropriate dimension, and M and $\widetilde{M}$ are positive sequences satisfying

(1) $M_0=1$,
(2) $M_k^2\le M_{k-1}M_{k+1}$,
(3) $M_k\le A^kM_jM_{k-j}$,
(4) $M_{k+1}^k\le A^kM_k^{k+1}$,
(5) $M_{k+1}/(kM_k)$ is increasing,
(6) $\sum_{k\ge0}M_k/M_{k+1}\le ApM_p/M_{p+1}$ for $p>0$,

where $A>0$ is a positive constant.

In the Gevrey case of interest to us, $M_k=k!^{\rho_1}$, $\widetilde{M}_k=k!^{\rho_2}$. For fixed $L_j>0$, the supremum in (D.1) defines a norm which equips a subspace of $\mathcal{C}^\infty_{M,\widetilde{M}}(X\times Y)$ with a Banach space structure. The space $\mathcal{C}^\infty_{M,\widetilde{M}}(X\times Y)$ is then the inductive limit of these spaces as $L=L_1=L_2\to\infty$, which identifies it a Silva space.

For $f\in\mathcal{C}^\infty_{M,\widetilde{M}}(X\times Y)$, and $z=(z_1,z_2)\in X\times Y$, $x\in X$ we define

$$(T_x^mf)(z):=\sum_{|\alpha|\le m}\frac{(\partial_x^\alpha f)(x,z_2)}{\alpha!}(z_1-x)^\alpha,\tag{D.2}$$

$$(R_x^mf)(z):=f(z)-(T_x^mf)(z).\tag{D.3}$$

To slightly generalise this notation, for a jet $f^{\alpha,\beta}$ of continuous functions, we write

$$(R_x^mf)_{\alpha,\beta}(z):=f^{\alpha,\beta}(z)-(T_x^{m-|\alpha|}f^{\alpha,\beta})(z).\tag{D.4}$$

We can now pose the question:

> Given a compact set $K\subset X$, under what conditions is it true that an arbitrary continuous jet $(f^{\alpha,\beta}):K\times Y\to\mathbb{R}$ is the jet of a function $\tilde{f}\in\mathcal{C}^\infty_{M,\widetilde{M}}(X\times Y)$?

We assume without loss of generality here that the set X is a full Euclidean space $\mathbb{R}^d$, rather than just an open subset thereof. This question is the anisotropic non-quasianalytic analogue of Whitney's extension theorem from classical analysis, which deals with the $\mathcal{C}^\infty$ case.

We begin by finding *necessary* conditions for the existence of such an extension, before proving that these conditions are indeed sufficient.

Proposition D.2. *Suppose $f \in \mathcal{C}^\infty_{M,\widetilde{M}}(X \times Y)$ with Gevrey constants L_1, L_2. Then there exists a constant A dependent only on the dimensions of X, Y and on M, $\widetilde{M}$ such that the jet $f^{\alpha,\beta} = \partial_z^{(\alpha,\beta)} f$ satisfies*

$$|f^{\alpha,\beta}| \le A L_1^{|\alpha|} L_2^{|\beta|} M_{|\alpha|} \widetilde{M}_{|\beta|}, \tag{D.5}$$

$$|(R_x^n f)_{k,l}(z)| \le A \tilde{L}_1^{n+1} M_{n+1} L_2^{|l|} \widetilde{M}_{|l|} \frac{|z_1 - x|^{n+1}}{(n+1)!} \tag{D.6}$$

for all nonnegative integers m, n and all multi-indices $|k| \le m$, $|l| \le n$, where $\tilde{L}_1 = C L_1$ with C dependent only on the dimension of X.

Proof. The first estimate (D.5) follows immediately from the definition of $\mathcal{C}^\infty_{M,\widetilde{M}}(X \times Y)$. We prove the second claim (D.6) by making use of the estimate (D.5) on the jet $f^{\alpha,\beta} = \partial_x^\alpha \partial_y^\beta f$ and Taylor expansion:

$$\begin{aligned} R_x^n f(z) &= \sum_{|\alpha|=n+1} \frac{n+1}{\alpha!} (z_1 - x)^\alpha \int_0^1 (1-t)^n f^{\alpha,0}(x + t(z_1 - x), z_2)\, dt \\ &\le \Big(\sup_{|\alpha|=n+1} \sup_{z \in X \times Y} |f^{\alpha,0}(z)| \Big) \sum_{|\alpha|=n+1} \left| \frac{(z_1 - x)^\alpha}{\alpha!} \right| \\ &\le \Big(\sup_{|\alpha|=n+1} \sup_{z \in X \times Y} |f^{\alpha,0}(z)| \Big) \frac{C^{n+1} |z_1 - x|^{n+1}}{(n+1)!}. \end{aligned} \tag{D.7}$$

Hence

$$|(R_x^n f)_{k,l}(z)| = |(R_x^{n-|k|} f)(z)| \le A \tilde{L}_1^{n+1} M_{n+1} L_2^{|l|} \widetilde{M}_{|l|} \frac{|z_1 - x|^{n+1}}{(n+1)!} \tag{D.8}$$

as required. □

Subsequently, for simplicity of notation, we omit the tilde in $\tilde{L}_1$ with the understanding that we are allowed to absorb constants that are dependent only on the dimensions of X, Y and on the sequences M, $\widetilde{M}$.

Theorem D.3. *Suppose $(f^{\alpha,\beta}) : K \times Y \to \mathbb{R}$ is a jet of continuous functions smooth in y that satisfies*

$$\partial_y^\gamma (f^{\alpha,\beta}) = f^{\alpha,\beta+\gamma}, \tag{D.9}$$

as well as the conditions (D.5) and (D.6) on $K \times Y$. Then there exists a function $f \in \mathcal{C}^\infty_{M,\widetilde{M}}(X \times Y)$ such that $\partial_x^{\alpha,\beta} f = f^{\alpha,\beta}$ on $K \times Y$.

Moreover, there exist constants C_0, C_1 dependent only on the dimensions of X and Y and the weight sequences M_k, $\widetilde{M}_k$ such that

$$\|f\|_{C_1 L_1, L_2} \le C_0 A. \tag{D.10}$$

Before proving Theorem D.3, we need to collect some lemmas, the proofs of which can be found in [Bruna 1980].

Proposition D.4. *Suppose $K \subset \mathbb{R}^d$ is compact. Then there exists a collection of closed cubes $\{Q_j\}_{j \in \mathbb{N}}$ with sides parallel to the axes such that*:

(i) $\mathbb{R}^d \setminus K = \bigcup_j Q_j$.

(ii) $\operatorname{int}(Q_j)$ *are disjoint.*

(iii) $\delta_j := \operatorname{diam}(Q_j) \le d_j := d(Q_j, K) \le 4\delta_j$.

(iv) *For* $0 < \lambda < \frac{1}{4}$, *we have* $d(z, K) \sim \delta_j$ *for* $z \in Q_j^* := (1+\lambda)Q_j$.

(v) *Each* Q_i^* *intersects at most* $D = (12)^{2d}$ *cubes* Q_j^*.

(vi) $\delta_i \sim \delta_j$ *if* $Q_i^* \cap Q_j^* \neq \varnothing$.

Proposition D.5. *For each* $\eta > 0$, *there exists a family of functions* $\phi_i \in \mathcal{C}_M^\infty(\mathbb{R}^d)$ *such that*:

(i) $0 \le \phi_i$.

(ii) $\operatorname{supp}(\phi_i) \subset Q_i^*$.

(iii) $\sum_i \phi_i(z) = 1$ *for* $z \in \mathbb{R}^d$.

(iv) $|\partial^\alpha \phi_i(z)| \le A h(B\eta d(z, K))\eta^{|\alpha|} M_{|\alpha|}$ *for* $z \in Q_i^*$, *where* $A, B > 0$ *are constants and*

$$h(t) := \sup_k \frac{k!}{t^k M_k}. \tag{D.11}$$

Proposition D.6. *Suppose* $T \in \mathcal{L}(E, F)$ *is a continuous linear surjection between Silva spaces. Then for any bounded set* $B \subset F$, *there exists a bounded set* $C \subset E$ *with* $T(C) = B$.

We also require an anisotropic version of Carleman's theorem, which is the special case of Theorem D.3 with $K = \{0\}$, and the Gevrey analogue of Borel's theorem from classical analysis.

Proposition D.7. *Let* $(g_\alpha)_{\alpha \in \mathbb{N}^d}$ *be a multisequence of functions in* $\mathcal{C}_{\widetilde{M}}^\infty(Y)$ *such that*

$$|\partial_y^l g_\alpha(y)| \le K L_1^{|\alpha|} L_2^{|l|} M_{|\alpha|} \widetilde{M}_{|l|} \tag{D.12}$$

for some constant $K > 0$.

Then there exists a function $f \in \mathcal{C}_{M,\widetilde{M}}^\infty(X \times Y)$ *such that* $g_\alpha(y) = \partial_x^\alpha f(0, y)$ *for all* $y \in Y$. *Moreover*, $\|f\|_{CL_1,L_2} \le AK$ *for some constants* $A, C > 0$ *independent of* f, L_1, *and* L_2.

Proof. We adapt the solution of [Petzsche 1988] of the classical Carleman problem to this setting. Key is that the assumptions on M imply that the hypotheses of [Petzsche 1988] are satisfied. Hence as in the proof of [Petzsche 1988, Theorem 2.1(ai)], we can construct compactly supported $\chi_p(x) \in \mathcal{C}_{M_p}^\infty(\mathbb{R})$ for each nonnegative integer p such that

$$\chi_p^{(k)}(0) = \delta(k, p) \tag{D.13}$$

and

$$\|\chi_p\|_{L(2+A^{-1})} \le \frac{1}{M_p}\left(\frac{Ae}{L}\right)^p \tag{D.14}$$

for some dimensional constant A and any $L > 0$. Hence we can define

$$\chi_\alpha(x) := \prod_{j=1}^d \chi_{\alpha_j}(x_j) \tag{D.15}$$

for $\alpha \in \mathbb{N}^d$ which satisfies

$$\chi_\alpha^{(\beta)}(0) = \delta(\beta, \alpha). \tag{D.16}$$

Moreover, we have the estimate

$$|\chi_\alpha^{(\beta)}| = \prod_{j=1}^{d} |\chi_{\alpha_j}^{\beta_j}| \le \prod_{j=1}^{d} \frac{1}{M_{\alpha_j}} \left(\frac{Ae}{L}\right)^{\alpha_j} (L(2+A^{-1}))^{\beta_j} M_{\beta_j}$$
$$\le \left(\frac{Aec(d,M)}{L}\right)^{|\alpha|} M_{|\alpha|}^{-1} (L(2+A^{-1}))^{|\beta|} M_{|\beta|}. \tag{D.17}$$

By taking $L = 2CL_1 = 2Aec(d,M)L_1$, we can estimate

$$|\partial_x^k \partial_y^l (\chi_\alpha(x) g_\alpha(y))| \le K((C/L)^{|\alpha|} M_{|\alpha|}^{-1} (L(2+A^{-1}))^{|k|} M_{|k|})(L_1^{|\alpha|} L_2^{|l|} M_{|\alpha|} \widetilde{M}_{|l|})$$
$$\le K \cdot 2^{-|\alpha|} (2CL_1(2+A^{-1}))^{|k|} L_2^{|l|} M_{|k|} \widetilde{M}_{|l|}, \tag{D.18}$$

where A, C, and K are constants independent of f, L_1, and L_2.

Hence we have that $\|\chi_\alpha(x) g_\alpha(y)\|_{2CL_1(2+A^{-1}),L_2} \le K \cdot 2^{-|\alpha|}$. It follows that

$$f(x,y) := \sum_{\alpha \in \mathbb{N}^d} \chi_\alpha(x) g_\alpha(y) \tag{D.19}$$

converges in the $\mathcal{C}^\infty_{M,\widetilde{M}}(X \times Y)$ sense, and satisfies $\partial_x^\alpha f(0,y) = g_\alpha(y)$ as required. □

Equipped with these tools, we are ready to prove Theorem D.3.

Proof of Theorem D.3. We begin by estimating the difference in Taylor expansions about different points in K. Using the identity

$$(T_x^n f)(z) - (T_y^n f)(z) = \sum_{|\alpha| \le n} \frac{(z_1 - x)^\alpha}{\alpha!} (R_y^n f)_{\alpha,0}(x, z_2), \tag{D.20}$$

we can estimate

$$\partial_z^{k,l}((T_x^n f)(z) - (T_y^n f)(z)) = \sum_{|\alpha| \le n-|k|} \frac{(z_1 - x)^\alpha}{\alpha!} (R_y^n f)_{k+\alpha,l}(x) \tag{D.21}$$

using the assumed estimate (D.6) for $(R_y^{m,n} f)_{k,l}$. This yields

$$|\partial_z^{k,l}((T_x^n f)(z) - (T_y^n f)(z))| \le AL_1^{n+1} M_{n+1} L_2^{|l|} \widetilde{M}_{|l|} \frac{(|z_1 - x| + |z_1 - y|)^{n-|k|+1}}{(n-|k|+1)!}. \tag{D.22}$$

We now invoke Proposition D.7. For $x \in X$ consider the map $T_x : \mathcal{C}^\infty_{M,\widetilde{M}}(X \times Y) \to G_x$ given by $(T_x f)_\alpha(y) := f^{\alpha,0}(x,y)$, where the space G_x consists of all multisequences of analytic functions $f_\alpha : Y \to \mathbb{R}$ satisfying $|f_\alpha| \le AL_1^{|\alpha|} L_2^{|\beta|} M_{|\alpha|} \widetilde{M}_{|\beta|}$ for some $A > 0$. From the assumed estimate (D.5) on $f^{\alpha,\beta}$, Proposition D.7 applies, and for each $x \in K$ we can find a function $f_x \in \mathcal{C}^\infty_{M,\widetilde{M}}(X \times Y)$ such that

$$\partial_z^{\alpha,\beta} f_x(x, z_2) = f^{\alpha,\beta}(x, z_2) \tag{D.23}$$

for each α, β. Moreover, the conclusion of Proposition D.7 implies that there exist constants $B = C_0 A$, $K_1 = C_1 L_1$, $K_2 = L_2 > 0$ such that the estimate

$$|(\partial_z^{\alpha,\beta} f_x)(z)| \le BK_1^{|\alpha|} K_2^{|\beta|} M_{|\alpha|} \widetilde{M}_{|\beta|} \tag{D.24}$$

holds uniformly, where the C_j depend only on the dimensions of X and Y and the weight sequences M_k, $\widetilde{M}_k$.

Hence we can bound

$$\partial_z^{k,l}(f_x(z) - (T_x^{m,n} f_x)(z)) = (R^{m,n} f_x)_{k,l}(z) \tag{D.25}$$

using the same calculation as in Proposition D.2. We obtain

$$\begin{aligned}|\partial_x^{k,l}(f_x(z) - (T_x^n f)(z))| &= |(R^n f_x)_{k,l}(z)| \\ &\le A(C_1L_1)^{n+1} M_{n+1} L_2^{|l|} \widetilde{M}_{|l|} \frac{|z_1 - x|^{n-|k|+1}}{(n-|k|+1)!}.\end{aligned} \tag{D.26}$$

The upshot of this estimate is that we can replace $T_x^n f$ and $T_y^n f$ in (D.22) with f_x and f_y respectively. That is, we have

$$|\partial_z^{k,l}(f_x(z) - f_y(z))| \le A(C_1L_1)^{n+1} M_{n+1} L_2^{|l|} \widetilde{M}_{|l|} \frac{(|z_1 - x| + |z_1 - y|)^{n-|k|+1}}{(n-|k|+1)!}. \tag{D.27}$$

We now fix k, l and vary $n \ge k$ in order to optimise the upper bound (D.27). By defining the quantity

$$h(t) := \sup_{k \ge 0} \frac{k!}{t^k M_k} \tag{D.28}$$

as in [Bruna 1980] we obtain

$$|\partial_z^{k,l}(f_x(z) - f_y(z))| \le A(C_1L_1)^{|k|} M_{|k|} L_2^{|l|} \widetilde{M}_{|l|} h((C_1L_1)(|z_1 - x_1| + |z_1 - y|))^{-1} \tag{D.29}$$

by using property (3) following (D.1).

The next step in the construction is to use Proposition D.5 to piece together the functions f_x using a $\mathcal{C}_M^\infty$ partition of unity subordinate to the cover arising from the decomposition of $X \setminus K$ by cubes in Proposition D.4. Taking the collection $\{Q_j\}_{j\in\mathbb{N}}$ of cubes in $X = \mathbb{R}^d$ constructed by Proposition D.4, we choose $x_j \in K$ such that $d(x_j, Q_j) = d(Q_j, K)$. Note that the conclusion of Proposition D.4 implies that

$$|z - x_j| \sim d(z, K) \tag{D.30}$$

for all $z \in Q_j^*$. Now taking ϕ_j as in Proposition D.5, we define

$$\tilde{f}(z) := \begin{cases} f(z) & \text{if } z_1 \in K, \\ \sum_i \phi_i(z_1) f_{x_j}(z) & \text{if } z_1 \in X \setminus K. \end{cases} \tag{D.31}$$

Note that since the partition of unity $\{\phi_j\}$ is locally finite, the function $\tilde{f}(z)$ is smooth in $(X \setminus K) \times Y$. It remains to check that $\tilde{f}$ is smooth elsewhere, and moreover that $\tilde{f} \in \mathcal{C}_{M,\widetilde{M}}^\infty(X \times Y)$. To this end, for $x \in K$ and $z_1 \in X \setminus K$, we estimate

$$\partial_z^{\alpha,\beta}(\tilde{f}(z) - f_x(z)) = \sum_{k \le \alpha} \binom{\alpha}{k} \sum_i (\partial^k \phi_i)(z_1)\, \partial_z^{\alpha-k,\beta}(f_{x_i}(z) - f_x(z)). \tag{D.32}$$

First we estimate the $k = 0$ term. If $z_1 \in \operatorname{spt}(\phi_i) = Q_i^*$, we have

$$d(z_1, x_i) \sim d(z_1, K) \le d(z_1, x) \tag{D.33}$$

and hence we have

$$\left|\sum_i \phi_i(z_1)\,\partial_z^{\alpha,\beta}(f_{x_i}(z)-f_x(z))\right| \le A(C_1L_1)^{|\alpha|}M_{|\alpha|}L_2^{|\beta|}\widetilde{M}_{|\beta|}h((C_1L_1)|z_1-x|)^{-1} \tag{D.34}$$

from (D.29).

We now estimate the terms with $|k|>0$. For $x\in X\setminus K$, we choose $\bar{x}\in K$ with $d(x,\bar{x})=d(x,K)$. Since $\sum_i \partial^k\phi_i=0$, we have

$$\sum_i(\partial^k\phi_i)(z_1)\,\partial_z^{\alpha-k,\beta}(f_{x_i}(z)-f_x(z))=\sum_i(\partial^k\phi_i)(z_1)\,\partial_z^{\alpha-k,\beta}(f_{x_i}(z)-f_{\bar{z}_1}(z)). \tag{D.35}$$

Now as before, we exploit the fact that $d(z_1,x_i)\sim d(z_1,K)$ to bound

$$|\partial_z^{\alpha-k,\beta}(f_{x_i}(z)-f_{\bar{z}_1}(z))|\le A(C_1L_1)^{|\alpha|-|k|}M_{|\alpha|-|k|}L_2^{|\beta|}\widetilde{M}_{|\beta|}h((C_1L_1)d(z_1,K))^{-1}. \tag{D.36}$$

Since $\log(M_j)$ is an increasing convex sequence with first term 0, it is also superadditive, and we have $M_{|k|}M_{|l|}\le M_{|k|+|l|}$. Hence for $|k|\ge 1$, we can use property (4) in Proposition D.5 to conclude that

$$\left|\sum_i(\partial^k\phi_i)(z_1)\,\partial_z^{\alpha-k,\beta}(f_{x_i}(z)-f_x(z))\right|\le AM_{|\alpha|}\widetilde{M}_{|\beta|}(C_1L_1)^{|\alpha|-|k|}L_2^{|\beta|}\eta^{|k|}\frac{h(B\eta d(z_1,K))}{h((C_1L_1)d(z_1,K))}, \tag{D.37}$$

where η remains to be chosen. Equation (15) from [Bruna 1980] implies the existence of a constant $c>0$ such that

$$\frac{h(t)}{h(ct)}\le\frac{A}{h(t)} \tag{D.38}$$

for some $A>0$. Hence we choose $\eta=C_1L_1/(cB)$ to arrive at the estimate

$$\left|\sum_i(\partial^k\phi_i)(z_1)\,\partial_z^{\alpha-k,\beta}(f_{x_i}(z)-f_x(z))\right|\le A(C_1L_1)^{|\alpha|-|k|}L_2^{|\beta|}M_{|\alpha|}\widetilde{M}_{|\beta|}\eta^{|k|}h((C_1L_1)|z_1-x|)^{-1}. \tag{D.39}$$

Combining (D.34) and (D.39), we arrive at

$$|\partial_z^{\alpha,\beta}(\tilde{f}(z)-f_x(z))|\le AL_2^{|\beta|}M_{|\alpha|}\widetilde{M}_{|\beta|}((C_1L_1)+\eta)^{|\alpha|}h((C_1L_1)|z_1-x|)^{-1} \tag{D.40}$$

for $z\in(X\setminus K)\times Y$.

The estimate (D.40) is key to proving $\tilde{f}\in\mathcal{C}^\infty(X\times Y)$ (and that the derivatives coincide with the those given by the jet $f^{\alpha,\beta}$), as well as the subsequent deduction of $\mathcal{C}^\infty_{M,\widetilde{M}}$ regularity. We write

$$\tilde{f}^{\alpha,\beta}(z):=\begin{cases}\partial_z^{\alpha,\beta}\tilde{f}(z) & \text{if } z_1\in X\setminus K,\\ f^{\alpha,\beta}(z) & \text{if } z_1\in K.\end{cases} \tag{D.41}$$

The smoothness of each $\tilde{f}^{\alpha,\beta}:X\times Y\to\mathbb{R}$ readily follows from the fact that each $f^{\alpha,\beta}:K\times Y\to\mathbb{R}$ is smooth in y, together with the estimate

$$|\tilde{f}^{\alpha,\beta}(z)-\partial_z^{\alpha,\beta}T_x^m f(z)|=o(|z_1-x|^{m-|\alpha|}). \tag{D.42}$$

For z with $z_1 \in K$, the estimate (D.42) comes immediately from (D.6) on $K \times Y$. Otherwise, it is a consequence of the estimate (D.40), the defining property (D.23) of the functions f_x, and the fact that the function $h(t)$ increases faster than any polynomial in t^{-1} as $t \to 0$.

Finally, we need to check $\mathcal{C}^{\infty}_{M,\widetilde{M}}$ regularity. That is, we need to verify the Gevrey estimate

$$\|f\|_{C_1 L_1, L_2} \leq C_0 A \tag{D.43}$$

for some constants C_0, C_1 dependent only on the dimensions of the spaces X and Y and the weight sequences M_k, $\widetilde{M}_k$. In light of (D.5), it only remains to prove (D.43) on $(X \setminus K) \times Y$, and by multiplication by a cutoff function we may assume $d(z_1, K)$ is bounded. Then, by applying (D.40) with $x = \bar{z}_1$ we can further reduce the problem to verifying (D.43) for f_x, uniformly in $x \in K$. However this was established earlier in (D.24). Hence, the proof is complete. □

References

[Arnold 1983] V. I. Arnold, *Geometrical methods in the theory of ordinary differential equations*, Grundlehren der Math. Wissenschaften **250**, Springer, 1983. MR Zbl

[Arnold 1989] V. I. Arnold, *Mathematical methods of classical mechanics*, 2nd ed., Grad. Texts in Math. **60**, Springer, 1989. MR Zbl

[Bruna 1980] J. Bruna, "An extension theorem of Whitney type for non-quasi-analytic classes of functions", *J. Lond. Math. Soc.* (2) **22**:3 (1980), 495–505. MR Zbl

[Colin de Verdière 1977] Y. Colin de Verdière, "Quasi-modes sur les variétés Riemanniennes", *Invent. Math.* **43**:1 (1977), 15–52. MR Zbl

[Colin de Verdière 1985] Y. Colin de Verdière, "Ergodicité et fonctions propres du laplacien", *Comm. Math. Phys.* **102**:3 (1985), 497–502. MR Zbl

[Duistermaat 1996] J. J. Duistermaat, *Fourier integral operators*, Progr. Math. **130**, Birkhäuser, Boston, 1996. MR Zbl

[Dyatlov and Guillarmou 2014] S. Dyatlov and C. Guillarmou, "Microlocal limits of plane waves and Eisenstein functions", *Ann. Sci. École Norm. Sup.* (4) **47**:2 (2014), 371–448. MR Zbl

[Gallavotti 1983] G. Gallavotti, "Perturbation theory for classical Hamiltonian systems", pp. 359–426 in *Scaling and self-similarity in physics* (Bures-sur-Yvette, France, 1981/1982), edited by J. Fröhlich, Progr. Phys. **7**, Birkhäuser, Boston, 1983. MR

[Gomes 2018] S. P. Gomes, "Percival's conjecture for the Bunimovich mushroom billiard", *Nonlinearity* **31**:9 (2018), 4108–4136. MR Zbl

[Gutkin 2009] B. Gutkin, "Note on converse quantum ergodicity", *Proc. Amer. Math. Soc.* **137**:8 (2009), 2795–2800. MR Zbl

[Hassell 2010] A. Hassell, "Ergodic billiards that are not quantum unique ergodic", *Ann. of Math.* (2) **171**:1 (2010), 605–618. MR Zbl

[Helffer et al. 1987] B. Helffer, A. Martinez, and D. Robert, "Ergodicité et limite semi-classique", *Comm. Math. Phys.* **109**:2 (1987), 313–326. MR Zbl

[Hou and Popov 2016] X. Hou and G. Popov, "Rigidity of reducibility of Gevrey quasi-periodic cocycles on $U(n)$", *Bull. Soc. Math. France* **144**:1 (2016), 1–52. MR Zbl

[Kato 1966] T. Kato, *Perturbation theory for linear operators*, Grundlehren der Math. Wissenschaften **132**, Springer, 1966. MR Zbl

[Knörrer 1980] H. Knörrer, "Geodesics on the ellipsoid", *Invent. Math.* **59**:2 (1980), 119–143. MR Zbl

[Kolmogorov 1954] A. N. Kolmogorov, "On conservation of conditionally periodic motions for a small change in Hamilton's function", *Dokl. Akad. Nauk SSSR (N.S.)* **98** (1954), 527–530. In Russian; translated as pp. 51–56 in *Stochastic behavior in* *classical and quantum Hamiltonian systems*, edited by G. Casati and J. Ford, Lecture Notes in Phys. **93**, Springer, 1979. MR Zbl

[Marklof and O'Keefe 2005] J. Marklof and S. O'Keefe, "Weyl's law and quantum ergodicity for maps with divided phase space", *Nonlinearity* **18**:1 (2005), 277–304. MR Zbl

[Moser 1966] J. Moser, "A rapidly convergent iteration method and non-linear partial differential equations, I", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (3) **20**:2 (1966), 265–315. MR Zbl

[Moser 1967] J. Moser, "Convergent series expansions for quasi-periodic motions", *Math. Ann.* **169** (1967), 136–176. MR Zbl

[Petzsche 1988] H.-J. Petzsche, "On E. Borel's theorem", *Math. Ann.* **282**:2 (1988), 299–313. MR Zbl

[Popov 2000a] G. Popov, "Invariant tori, effective stability, and quasimodes with exponentially small error terms, I: Birkhoff normal forms", *Ann. Henri Poincaré* **1**:2 (2000), 223–248. MR Zbl

[Popov 2000b] G. Popov, "Invariant tori, effective stability, and quasimodes with exponentially small error terms, II: Quantum Birkhoff normal forms", *Ann. Henri Poincaré* **1**:2 (2000), 249–279. MR Zbl

[Popov 2004a] G. Popov, "KAM theorem and quasimodes for Gevrey Hamiltonians", *Mat. Contemp.* **26** (2004), 87–107. MR Zbl

[Popov 2004b] G. Popov, "KAM theorem for Gevrey Hamiltonians", *Ergodic Theory Dynam. Systems* **24**:5 (2004), 1753–1786. MR Zbl

[Pöschel 2001] J. Pöschel, "A lecture on the classical KAM theorem", pp. 707–732 in *Smooth ergodic theory and its applications* (Seattle, WA, 1999), edited by A. Katok et al., Proc. Sympos. Pure Math. **69**, Amer. Math. Soc., Providence, RI, 2001. MR Zbl

[Shnirelman 1974] A. I. Shnirelman, "Ergodic properties of eigenfunctions", *Uspekhi Mat. Nauk* **29**:6(180) (1974), 181–182. In Russian. MR Zbl

[Toth and Zelditch 2003] J. A. Toth and S. Zelditch, "Norms of modes and quasi-modes revisited", pp. 435–458 in *Harmonic analysis at Mount Holyoke* (South Hadley, MA, 2001), edited by W. Beckner et al., Contemp. Math. **320**, Amer. Math. Soc., Providence, RI, 2003. MR Zbl

[Zelditch 1987] S. Zelditch, "Uniform distribution of eigenfunctions on compact hyperbolic surfaces", *Duke Math. J.* **55**:4 (1987), 919–941. MR Zbl

[Zelditch 2004] S. Zelditch, "Note on quantum unique ergodicity", *Proc. Amer. Math. Soc.* **132**:6 (2004), 1869–1872. MR Zbl

[Zworski 2012] M. Zworski, *Semiclassical analysis*, Grad. Stud. in Math. **138**, Amer. Math. Soc., Providence, RI, 2012. MR Zbl

Received 8 Mar 2020. Revised 25 Feb 2021. Accepted 10 Jun 2021.

SEÁN GOMES: sean.p.gomes@gmail.com
Department of Mathematics, Northwestern University, Chicago, IL, United States

ANALYSIS AND PDE
Vol. 16 (2023), No. 1, pp. 173–232
DOI: 10.2140/apde.2023.16.173

STRICHARTZ ESTIMATES FOR MIXED HOMOGENEOUS SURFACES IN THREE DIMENSIONS

LJUDEVIT PALLE

We obtain sharp mixed-norm Strichartz estimates associated to mixed homogeneous surfaces in $\mathbb{R}^3$. Cases with and without a damping factor are both considered. In the case when a damping factor is considered our results yield a wide generalization of a result of Carbery, Kenig, and Ziesler for homogeneous polynomial surfaces in $\mathbb{R}^3$. The approach we use is to first classify all possible singularities locally, after which one can tackle the problem by appropriately modifying the methods from a paper of Ginibre and Velo, and by using the recently developed methods by Ikromov and Müller.

1. Introduction

Let us fix a pair $\alpha = (\alpha_1, \alpha_2) \in (0, \infty)^2$, define $|\alpha| := \alpha_1 + \alpha_2$, and introduce its associated α-mixed homogeneous dilations in $\mathbb{R}^2$ by

$$\delta_t(x_1, x_2) = (t^{\alpha_1} x_1, t^{\alpha_2} x_2), \quad t > 0.$$

The main goal of this article is to study Strichartz estimates for a fixed mixed homogeneous surface S, i.e., a surface given as the graph of a fixed smooth function $\phi : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}$ which is α-mixed homogeneous of degree ρ:

$$\phi \circ \delta_t(x_1, x_2) = t^\rho \phi(x_1, x_2), \quad t > 0. \tag{1-1}$$

We may and shall assume without loss of generality that $\rho \in \{-1, 0, 1\}$. Both α and ρ shall be fixed throughout the article. Note that when $\rho = -1$ the function ϕ has a singularity at the origin.

The author was supported by Deutsche Forschungsgemeinschaft, project number 237750060, *Fragen der Harmonischen Analysis im Zusammenhang mit Hyperflächen.*
MSC2020: 35B45, 42B15.
Keywords: Strichartz estimates, Fourier restriction, oscillatory integrals, classification of singularities, mixed homogeneous, mixed-norm.

As is well known, Strichartz estimates are directly related to Fourier restriction estimates and we are in particular interested in the mixed-norm estimate

$$\|\hat{f}\|_{L^2(\mathrm{d}\mu)} \le C\|f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})}, \qquad f \in \mathcal{S}(\mathbb{R}^3), \tag{1-2}$$

where μ is the surface measure

$$\langle \mu, f\rangle = \int_{\mathbb{R}^2\setminus\{0\}} f(x_1,x_2,\phi(x_1,x_2))\,\mathcal{W}(x_1,x_2)\,\mathrm{d}x \tag{1-3}$$

and $p=(p_1,p_3)\in(1,2)^2$. Note that we skip the p_2-exponent which corresponds to the integration in the x_2-variable — here we consider the case $p_1=p_2$, i.e., we have one exponent $p_1=p_2$ in the "tangential" direction and another, namely p_3, in the "normal" direction to the surface S at $(0,0,\phi(0,0))$ (this will be formally true only when ϕ is smooth at the origin).

The weight $\mathcal{W}\ge 0$ is added in order to ensure that the measure has a scaling invariance which will enable us to reduce global estimates to local ones by a Littlewood–Paley argument. We take $\mathcal{W}$ to be α-mixed homogeneous of degree 2ϑ and consider two particular cases. The function $\mathcal{W}$ will be either equal to

$$|x|_\alpha^{2\vartheta} = (|x_1|^{1/\alpha_1}+|x_2|^{1/\alpha_2})^{2\vartheta} \tag{1-4}$$

or equal to the Hessian determinant of ϕ (denoted by $\mathcal{H}_\phi$) raised to the power $|\cdot|^\sigma$, $\sigma\in\left[0,\frac{1}{2}\right)$, i.e.,

$$|\mathcal{H}_\phi(x)|^\sigma = \left|\det\begin{bmatrix}\partial_{x_1}^2\phi & \partial_{x_1}\partial_{x_2}\phi\\ \partial_{x_1}\partial_{x_2}\phi & \partial_{x_2}^2\phi\end{bmatrix}\right|^\sigma. \tag{1-5}$$

The first weight (1-4) is of interest as a type of mixed homogeneous Sobolev weight, while the second one (1-5) was considered originally in [Sjölin 1974] and turns out to be a natural choice when studying Fourier restriction estimates for surfaces with vanishing Gaussian curvature. One can easily show that the Hessian determinant of ϕ is α-mixed homogeneous of degree $2(\rho-|\alpha|)$, and so in the case when $\mathcal{W}$ equals (1-5) the relation between ϑ and σ is $\vartheta=\sigma(\rho-|\alpha|)$. We shall later determine ϑ in Section 2A (and in particular in Proposition 2.1) so that the Fourier restriction estimate for μ is invariant under scaling. This choice depends in general on $p=(p_1,p_3)$.

Oscillatory integrals, Fourier restriction estimates, and other problems related to homogeneous and mixed homogeneous surfaces have been previously studied in works such as [Dendrinos and Zimmermann 2019; Schwend 2020; Greenblatt 2018; Ikromov et al. 2005; Ikromov and Müller 2011; Iosevich and Sawyer 1996; Ferreyra et al. 2004; Ferreyra and Urciuolo 2009; Carbery et al. 2013].

The case of general L^p-L^2 Fourier restriction in $\mathbb{R}^3$ with respect to the Euclidean measure was recently solved in [Ikromov and Müller 2016] for a wide class of smooth surfaces in $\mathbb{R}^3$, including all the analytic ones. Mixed-norm estimates were shown in [Palle 2021] for surfaces given as graphs of functions ϕ in adapted coordinates and also for analytic functions ϕ satisfying $h_{\mathrm{lin}}(\phi)<2$ (see below for the definition of linear height $h_{\mathrm{lin}}(\phi)$).

In [Carbery et al. 2013] Carbery, Kenig, and Ziesler considered the case with the weight (1-5) for "isotropically" homogeneous (i.e., when $\alpha_1=\alpha_2$) polynomials ϕ. Since the weight (1-5) has roots at the degenerate points, the estimate (1-2) holds for a wider range of exponents compared to the case when the weight (1-4) is used. As already mentioned, the use of this so-called mitigating or damping factor goes back to [Sjölin 1974] (see also [Cowling et al. 1990; Drury 1990; Kenig et al. 1991]). Its naturalness stems from the fact that it is equiaffine invariant as is the Fourier transformation. In fact, the mitigating factor can be expressed in a parametrization-independent way through the use of so-called affine fundamental forms (see, e.g., [Su 1983; Nomizu and Sasaki 1994]). When one uses the above damping factor (1-5) one can even obtain estimates for certain classes of flat surfaces [Carbery and Ziesler 2002; Abi-Khuzam and Shayya 2006; Carbery et al. 2007]. On the other hand, weak-type $L^{4/3}$-$L^{4(n-1)/(n+1)}$ estimates were obtained in [Oberlin 2012] for a wide class of surfaces having a bounded generic multiplicity (see also [Oberlin 2004]). In the three-dimensional case ($n = 3$) they correspond to precisely the Tomas–Stein range, but otherwise are a strict subset of it. Let us also mention a recent result of [Gressman 2016] where he obtained decay estimates for damping oscillatory integrals for a certain class of singularities.

In this article we shall first classify the possible local singularities for mixed homogeneous surfaces (see Proposition 1.4 below) and then either apply the Fourier restriction estimates obtained in [Ikromov and Müller 2016; Palle 2021] or use the techniques from these articles, and also from [Ginibre and Velo 1992] (see also [Keel and Tao 1998]), to obtain sharp estimates. In particular, we obtain a wide generalization of the Fourier restriction estimate in [Carbery et al. 2013] with methods which are more elementary and avoiding any use of results from algebraic topology or algebraic geometry. Namely, in [Carbery et al. 2013] a result of [Milnor 1964] on Betti numbers is used in order to control the number of connected components of a set given by polynomial inequalities.

In order to state the main results of this paper (namely, Theorem 1.1, Theorem 1.2, Proposition 1.4, and Corollary 1.5) we first recall certain concepts and introduce a few conditions. Recall that the *Taylor support* of a smooth function $\varphi : \Omega \subseteq \mathbb{R}^2 \to \mathbb{R}$ at $P \in \Omega$ is defined as the set $\mathcal{T}(\varphi, P) := \{\tau \in \mathbb{N}_0^2 : \partial^\tau \varphi(P) \neq 0\}$. We call φ a function of *finite type* at P if its Taylor support at P is nonempty. If φ is of finite type at P, then one defines its *Newton polyhedron* $\mathcal{N}(\varphi, P)$ at P as the convex hull of the union of sets $\{(t_1, t_2) \in \mathbb{R}^2 : t_1 \geq \tau_1,\ t_2 \geq \tau_2\}$, where $\tau = (\tau_1, \tau_2)$ goes over the Taylor support of φ at P.

We can now recall the definitions of some very important quantities from the theory of oscillatory integrals which go back to V. I. Arnold and A. N. Varchenko (see, e.g., [Varchenko 1976]). Let us assume for a function φ of finite type at P that $\varphi(P) = 0$ and $\nabla\varphi(P) = 0$. If this is not the case we simply subtract the constant and linear terms of the Taylor series of φ at P. The *Newton distance* $d(\varphi, P)$ of φ at P is then defined as the minimum of the set $\{t \in \mathbb{R} : (t, t) \in \mathcal{N}(\varphi, P)\}$. The face (i.e., a vertex or an edge) where the line $\{(t, t) : t \in \mathbb{R}\}$ intersects the Newton polyhedron $\mathcal{N}(\varphi, P)$ is called the *principal face* and it is denoted by $\pi(\varphi, P)$. Note that $d(\varphi, P) \geq 1$. The *Newton height* $h(\varphi, P)$ of φ at P is defined as the supremum of the set $\{d(\varphi \circ \Phi, P) : \Phi \text{ is a local diffeomorphism at } P\}$. We define the *linear height* $h_{\mathrm{lin}}(\varphi, P)$ analogously — the only difference is that one considers linear coordinate changes centered at P instead of local diffeomorphisms. Note that $h(\varphi, P) \geq h_{\mathrm{lin}}(\varphi, P) \geq d(\varphi, P)$. One says that φ is *adapted* at P if $d(\varphi, P) = h(\varphi, P)$ and that it is *linearly adapted* at P if $d(\varphi, P) = h_{\mathrm{lin}}(\varphi, P)$. Similarly, one says that φ is *adapted* in the Φ coordinates if $d(\varphi \circ \Phi, P) = h(\varphi, P)$ and one defines what it means to be *linearly adapted* in the Φ coordinates analogously. The existence of a coordinate system in which an analytic function is adapted was shown in [Varchenko 1976]. This was generalized to smooth functions of finite type in [Ikromov and Müller 2011]. For the existence of a linear coordinate change in which a function is linearly adapted see [Ikromov and Müller 2016].

Finally, for a function φ of finite type at P satisfying $\varphi(P)=0$ and $\nabla\varphi(P)=0$ we recall the definition of *Varchenko's exponent*, denoted by $\nu(\varphi, P)$. It is defined to be 1 if $h(\varphi, P)\geq 2$ and if there exists a coordinate change Φ in which ϕ is adapted and so that the bisectrix $\{(t,t):t\in\mathbb{R}\}$ intersects the Newton polyhedron $\mathcal{N}(\varphi\circ\Phi, P)$ at a vertex. Otherwise one defines $\nu(\varphi, P):=0$.

The relation to oscillatory integrals is as follows. If one is given a smooth amplitude a localized at P, then the decay rate of the oscillatory integral $\int a(x)e^{i\lambda\varphi(x)}\,dx$ is $\lambda^{-1/h(\varphi,P)}(\log\lambda)^{\nu(\varphi,P)}$ for large λ. This also holds when one considers small linear perturbations of φ.

Let us mention that one often translates P to 0, in which case one uses the notation $\mathcal{T}(\varphi)$, $d(\varphi)$, $\nu(\varphi)$, etc., and it is implicitly understood that everything is considered at the origin.

In this article we shall consider either of the following two conditions on our fixed α-mixed homogeneous function ϕ:

(H1) At any given point $(x_1,x_2)\neq(0,0)$ where the Hessian determinant of ϕ vanishes at least one of the mappings $t\mapsto\partial_1^2\phi(t,x_2)$ or $t\mapsto\partial_2^2\phi(x_1,t)$ is of finite type at $t=x_1$ (resp. $t=x_2$), i.e., at least one of them or their derivatives is nonzero when evaluated at the respective points.

(H2) The Hessian determinant $\mathcal{H}_\phi$ is not flat at any point $x\neq 0$.

It actually suffices to check the conditions only at points (x_1,x_2) in, say, a unit circle by homogeneity. Furthermore, we remark that the condition (H2) is stronger than the condition (H1) (this follows from the calculations in Section 3B below).

Let us now introduce a further condition and two new quantities. For a point $v\in\mathbb{R}^2\setminus\{0\}$ let us define the function

$$\phi_v(x):=\phi(x+v)-\phi(v)-x\cdot\nabla\phi(v).$$

Then we shall often consider whether the following condition is satisfied at v:

(LA) There is a linear coordinate change which is adapted to ϕ_v at the origin.

Compare with the negation of this condition in [Ikromov and Müller 2016, Section 1.2]. Note that the (LA) condition is not the same as linear adaptedness of ϕ_v at 0.

As mentioned, the linear height of ϕ_v and the Newton height of ϕ_v are respectively denoted by $h_{\text{lin}}(\phi, v)$ and $h(\phi, v)$. We define the *global linear height* $h_{\text{lin}}(\phi)$ and the *global Newton height* $h(\phi)$ by the respective expressions

$$h_{\text{lin}}(\phi)=\sup_{v\in\mathbb{S}^1}h_{\text{lin}}(\phi,v),\quad h(\phi)=\sup_{v\in\mathbb{S}^1}h(\phi,v). \tag{1-6}$$

It will be clear from Section 3 that $h_{\text{lin}}(\phi, v)$ and $h(\phi, v)$ do not change along the *homogeneity curve through* v defined as the curve

$$t\mapsto(t^{\alpha_1}v_1,t^{\alpha_2}v_2),\quad t>0,$$

and therefore in the above definitions of global linear height and global Newton height one could have taken the supremum over the set $\mathbb{R}^2\setminus\{0\}$ too.

Theorem 1.1. *Let ϕ be mixed homogeneous satisfying condition (H2). Let μ be the measure defined as in (1-3) with $\mathcal{W}(x) = |\mathcal{H}_\phi(x)|^\sigma$ for some fixed $\sigma \geq 0$. If $\sigma \in \left[0, \frac{1}{3}\right]$, then the Fourier restriction estimate (1-2) holds true for*

$$\left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1}{2} - \sigma, \sigma\right).$$

If (LA) is satisfied at all points $v \neq 0$, then the estimate holds true even if $\sigma \in \left[0, \frac{1}{2}\right)$. In particular, if $\alpha_1 = \alpha_2$, then (LA) is automatically satisfied at all points $v \neq 0$ and the estimate holds true for any $\sigma \in \left[0, \frac{1}{2}\right)$.

Several comments are in order. Firstly, precise conditions for when the (LA) condition is satisfied at $v \neq 0$ can be checked by using the normal-form tables in Section 3 (note that in the Proposition 1.4 below, where the normal forms are listed, only the normal form (vi) is not in adapted coordinates). That one is restricted to $0 \leq \sigma \leq \frac{1}{3}$ in the case when (LA) is not satisfied is a consequence of a Knapp-type example, as we shall show in Section 4F1. That the result in the above theorem is sharp is well known — as soon as one knows that the Hessian determinant of ϕ does not vanish identically we can apply the classical Knapp example to a point where the Hessian does not vanish which then yields the necessary condition

$$\frac{1}{p_1'} + \frac{1}{p_3'} \leq \frac{1}{2}.$$

Secondly, in the case when $\rho = 1 = |\alpha|$, one can extend the above estimate to the range where

$$\frac{1}{p_1'} + \frac{1}{p_3'} = \frac{1}{2}, \qquad \frac{1}{p_3'} \leq \sigma.$$

The reason for this is that $\rho = 1 = |\alpha|$ implies that the weight $\mathcal{W} = |\mathcal{H}_\phi|^\sigma$ (and the Hessian determinant) are α-mixed homogeneous of degree 0, and hence bounded on $\mathbb{R}^2$, and so the estimate for $(p_1, p_3) = (2, 1)$ follows trivially by Plancherel.

Finally, let us mention that the most interesting part of the proof of the above theorem is the proof of Fourier restriction for the normal form (v) from Proposition 1.4, which is to be found in Section 4E. There we need to estimate the Fourier transform of a certain measure, and for this we perform a natural decomposition of this measure. What is remarkable is that at the critical frequencies one initially has an infinite number of pieces which are not summable absolutely, but, after a delicate analysis, only $\mathcal{O}(1)$ decomposition pieces turn out to have a "bad" estimate. Interestingly, a similar thing happens in the much easier case of normal form (iv).

In the case of the other weight (which has no roots away from the origin) we have:

Theorem 1.2. *Let ϕ be mixed homogeneous satisfying condition (H1). Let μ be the measure defined as in (1-3) with $\mathcal{W}(x) = |x|_\alpha^{2\vartheta}$. If the exponents $(p_1, p_3) \in (1, 2)^2$ and $\vartheta \in \mathbb{R}$ satisfy* (see Figure 1)

$$\frac{1}{p_1'} + \frac{h_{\text{lin}}(\phi)}{p_3'} \leq \frac{1}{2}, \qquad \frac{1}{p_3'} \leq \frac{1}{2h(\phi)}, \qquad \vartheta = \frac{|\alpha|}{p_1'} + \frac{\rho}{p_3'} - \frac{|\alpha|}{2},$$

then the Fourier restriction estimate (1-2) holds true.

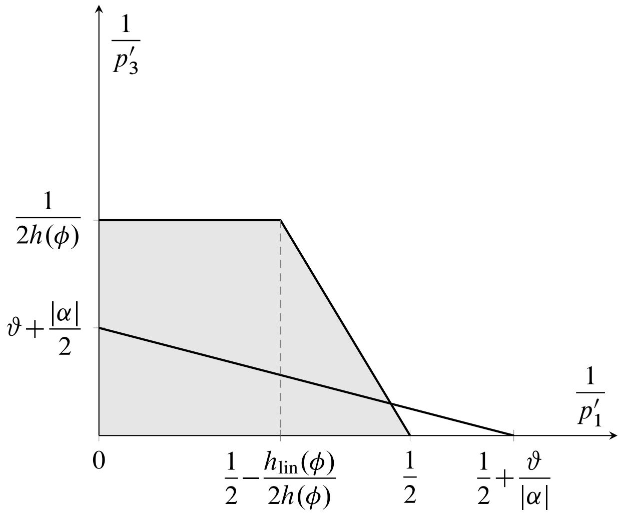

Figure 1. The Riesz diagram for the range of exponents given in Theorem 1.2. The line given by ϑ is drawn for the case when $\rho = 1$, $\vartheta > 0$, and both $|\alpha|$ and ϑ are small.

We remark that the quantity ϑ in the above theorem is allowed to be negative. This theorem is sharp since the corresponding local estimates are sharp — this was shown in [Palle 2021]. We discuss this in more detail at the beginning of Section 5.

As a special case of Theorem 1.1 we obtain:

Corollary 1.3. *Let ϕ be any mixed homogeneous polynomial in $\mathbb{R}^2$ and let μ be the measure defined as in (1-3) with $\mathcal{W}(x) = |\mathcal{H}_\phi(x)|^{1/4}$. Then the Fourier restriction estimate (1-2) holds true for $p_1' = p_3' = 4$.*

In the case of the above corollary we note that the Hessian determinant can either vanish identically, or it does not vanish to infinite order anywhere, since it is necessarily a nonzero mixed homogeneous polynomial. But the case when the Hessian determinant vanishes identically is trivial, so we are indeed within the scope of Theorem 1.1.

When one considers "isotropically" homogeneous polynomials (i.e., when $\alpha_1 = \alpha_2$), Corollary 1.3 recovers the main result of [Carbery et al. 2013]. The strategy of proof in that work was to first perform certain decompositions of the surface measure in order to get appropriate control over the size of $\nabla\phi$ and the Hessian determinant $\mathcal{H}_\phi$, after which one applies an L^4 argument, as the $L^{4/3}(\mathbb{R}^3) \to L^2(\mathrm{d}\mu)$ Fourier restriction estimate is equivalent to the $L^2(\mathrm{d}\mu) \to L^4(\mathbb{R}^3)$ extension estimate.

Our proofs of Theorems 1.1 and 1.2 are based on the following intermediary result:

Proposition 1.4. *Let $v \in \mathbb{R}^2 \setminus \{0\}$, let ϕ be as above α-mixed homogeneous of degree ρ, and let us assume that it satisfies condition (H1) and that its Hessian determinant vanishes at v. Then after a linear transformation of coordinates the function $\phi_v(x) := \phi(x+v) - \phi(v) - x \cdot \nabla\phi(v)$ and its Hessian determinant $\mathcal{H}_{\phi_v}$ assume precisely one of the normal forms in Table 1. In all the cases the appearing functions are smooth and do not vanish at the origin, i.e., $r(0), r_0(0), r_1(0), r_2(0), q(0), \psi(0) \neq 0$, except for the function φ which is flat at the origin.*

case	normal form	additional conditions
(i)	$\phi_v(x) = x_2^k r(x) + \varphi(x)$, $\mathcal{H}_{\phi_v}(x) = x_2^{\tilde{k}+2k-2} r_0(x)$ or $\mathcal{H}_{\phi_v}$ flat at 0	$k \geq 2,\ \tilde{k} \geq 0$
(ii)	$\phi_v(x) = x_1^2 q(x_1) + x_2^k r(x)$, $\mathcal{H}_{\phi_v}(x) = x_2^{k-2} r_0(x)$	$k \geq 3$
(iii)	$\phi_v(x) = x_1^2 r_1(x) + x_2^k r_2(x)$, $\mathcal{H}_{\phi_v}(x) = x_2^{k-2} r_0(x)$	$k \geq 3$, $\partial_2^j r_1(0) = c(\phi, v)\, j\, \partial_2^{j-1} r_1(0)$ for $c(\phi, v) \neq 0,\ j = 1, \dots, k-1$
(iv)	$\phi_v(x) = x_1^2 q(x_1) + (x_2 - x_1^2 \psi(x_1))^k r(x)$, $\mathcal{H}_{\phi_v}(x) = (x_2 - x_1^2 \psi(x_1))^{k-2} r_0(x)$	$k \geq 3$
(v)	$\phi_v(x) = x_1^2 r_1(x) + (x_2 - x_1^2 \psi(x_1))^k r_2(x)$, $\mathcal{H}_{\phi_v}(x) = (x_2 - x_1^2 \psi(x_1))^{k-2} r_0(x)$	$k \geq 3$, $\partial_2^j r_1(0) = c(\phi, v)\, j\, \partial_2^{j-1} r_1(0)$ for $c(\phi, v) \neq 0,\ j = 1, \dots, k-1$
(vi)	$\phi_v(x) = (x_2 - x_1^2 \psi(x_1))^k r(x)$, $\mathcal{H}_{\phi_v}(x) = (x_2 - x_1^2 \psi(x_1))^{2k-3} r_0(x)$	$k \geq 2$

Table 1. Normal forms for Proposition 1.4.

In the case of normal form (i) *one additionally knows that if the Hessian determinant $\mathcal{H}_{\phi_v}$ is not flat at the origin, then φ vanishes identically. In particular, if condition* (H2) *is satisfied, then the function φ in case* (i) *always vanishes identically and the Hessian determinant is nowhere flat. In the case when $\alpha_1 = \alpha_2$ the functions ϕ_v and $\mathcal{H}_{\phi_v}$ can only take the forms* (i) *or* (ii). *Finally, the root of the function $x \mapsto x_2 - x_1^2 \psi(x_1)$ corresponds to the homogeneity curve through v, though in the coordinate system in which the normal form is given.*

In cases (i) and (ii) one has further subcases (see Section 3A) of a technical nature, so we left them out of the above proposition. We also note that only in case (vi) the function ϕ_v is not in adapted coordinates (and the adapted coordinates can be achieved only through a nonlinear transformation such as $(x_1, x_2) \mapsto (x_1, x_2 + x_1^2 \psi(x_1))$), but it is linearly adapted.

The idea to apply Fourier restriction estimates to obtain a priori estimate for PDEs goes back to [Strichartz 1977]. In our case one can apply the above results to obtain Strichartz estimates for the nonhomogeneous initial problem

$$\begin{cases} (\partial_t - i\phi(D))u(x,t) = F(x,t), & (x,t) \in \mathbb{R}^2 \times (0,\infty), \\ \qquad\qquad u(x,0) = G(x), & x \in \mathbb{R}^2, \end{cases}$$

where $F \in \mathcal{S}(\mathbb{R}^3)$, $G \in \mathcal{S}(\mathbb{R}^2)$. Namely, by an application of the Christ–Kiselev lemma [2001] one gets the following result:

Corollary 1.5. *Let ϕ, $\mathcal{W}$, and $(p_1, p_3) \in (1,2)^2$ be either as in Theorem 1.1 or 1.2, and let us furthermore assume that $\rho \in \{0,1\}$. Then for the above nonhomogeneous PDE one has the a priori estimate*

$$\|u\|_{L_t^{p_3'}(L_{(x_1,x_2)}^{p_1'})} \le C_1 \|\mathcal{W}^{-1/2}\mathcal{F}G\|_{L^2(\mathbb{R}^2)} + C_2 \|\mathcal{F}_{(x_1,x_2)}^{-1}(\mathcal{W}^{-1}\mathcal{F}_{(x_1,x_2)}F)\|_{L_t^{p_3}(L_{(x_1,x_2)}^{p_1})},$$

where $\mathcal{F}_{(x_1,x_2)}$ is the partial Fourier transformation in the $x = (x_1, x_2)$-direction.

In the case when $\mathcal{W}$ is the function $|\cdot|_\alpha^{2\vartheta}$ the norms on the right-hand side are a type of homogeneous anisotropic Sobolev norms [Triebel 2006, Chapter 5] (in particular, note that $\|\mathcal{W}^{-1/2}\mathcal{F}G\|_{L^2(\mathbb{R}^2)} = \|\mathcal{F}^{-1}\mathcal{W}^{-1/2}\mathcal{F}G\|_{L^2(\mathbb{R}^2)}$).

Since the procedure of how to obtain the corresponding Strichartz estimate from a Fourier restriction estimate is mostly standard we have deferred the sketch of the proof of Corollary 1.5 to the Appendix.

The article is structured in the following way. In Section 2 we first perform some elementary reductions. Since the proofs of Theorems 1.1 and 1.2 are essentially based on Proposition 1.4, we first prove this proposition (and even obtain slightly more precise results) in Section 3. Subsequently we prove Theorems 1.1 and 1.2 in Sections 4 and 5 respectively. In the Appendix we then give a sketch of the proof of Corollary 1.5.

In this paper we use the symbols $\sim$, $\lesssim$, $\gtrsim$, $\ll$, $\gg$ with the following meanings. If two nonnegative quantities A and B are given, then by $A \ll B$ we mean that there exists a sufficiently small positive constant c such that $A \le cB$, and by $A \lesssim B$ we mean that there exists a (possibly large) positive constant C such that $A \le CB$. The relation $A \sim B$ means that there exist positive constants $C_1 \le C_2$ such that $C_1 A \le B \le C_2 A$ is satisfied. Relations $A \gg B$ and $A \gtrsim B$ are defined analogously. Sometimes the implicit constants c, C, C_1, and C_2 depend on certain parameters p, and in order to emphasize this dependence we shall write for example $\lesssim_p$, $\sim_p$, and so on.

We also use the symbols χ_0, χ_1, r, and q generically in the following way. We require χ_0 to be supported in a neighborhood of the origin and identically equal to 1 near the origin. On the other hand, we require χ_1 to be supported away from the origin and identically equal to 1 on an open neighborhood of $\pm 1 \in \mathbb{R}$. Sometimes, when several χ_0 or χ_1 appear within the same formula, they may designate different functions. The functions r and q (also used with subscripts and tildes) shall be used generically as smooth functions which are nonvanishing at the origin, where the function q shall denote a function of one variable, whereas the function r shall denote a function which may generally depend on two variables. Occasionally both of them can also be flat at the origin, in which case we state this explicitly.

2. Preliminary reductions

2A. ***Rescaling and reduction to local estimates.*** As mentioned, the measure we consider is

$$\langle \mu, f \rangle = \int_{\mathbb{R}^2 \setminus \{0\}} f(x_1, x_2, \phi(x_1, x_2))\, \mathcal{W}(x_1, x_2)\, \mathrm{d}x,$$

where $\mathcal{W}$ is nonnegative, continuous on $\mathbb{R}^2 \setminus \{0\}$, and α-mixed homogeneous of degree 2ϑ. In this subsection we determine the degree of homogeneity 2ϑ so that the global Fourier restriction estimate (1-2) becomes equivalent to the local one. By this we mean the following. Let us take a partition of unity $(\eta_j)_{j\in\mathbb{Z}}$ in $\mathbb{R}^2 \setminus \{0\}$,

$$\sum_{j\in\mathbb{Z}} \eta_j(x) = 1, \quad x \neq 0, \tag{2-1}$$

such that $\eta_j = \eta \circ \delta_{2^{-j}}$ for some $\eta = \eta_0 \in C_c^\infty(\mathbb{R}^2)$ supported away from the origin. Let us consider the measures

$$\langle \mu_j, f \rangle := \int_{\mathbb{R}^2} f(x, \phi(x))\, \eta_j(x)\, \mathcal{W}(x)\, \mathrm{d}x, \tag{2-2}$$

which now satisfy $\mu = \sum_{j\in\mathbb{Z}} \mu_j$, and let us furthermore assume that we have the local estimate for some $j_0 \in \mathbb{Z}$:

$$\|\hat{f}\|_{L^2(\mathrm{d}\mu_{j_0})} \le C \|f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})}.$$

We want to determine the degree of homogeneity of $\mathcal{W}$ so that the Fourier restriction estimate is invariant under the dilations δ_t, i.e., that we have

$$\|\hat{f}\|_{L^2(\mathrm{d}\mu_j)} \le C \|f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})} \tag{2-3}$$

for all $j \in \mathbb{Z}$ whenever the estimate is true for some $j_0 \in \mathbb{Z}$. In this case, and if $(p_1, p_3) \in (1, 2]^2$, a standard Littlewood–Paley argument (presented below) will then yield

$$\|\hat{f}\|_{L^2(\mathrm{d}\mu)} \le C \|f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})}.$$

To summarize, we have:

Proposition 2.1. *Let $\mathcal{W}$ be α-mixed homogeneous of degree 2ϑ, not identically zero, and continuous on $\mathbb{R}^2 \setminus \{0\}$, let μ be defined as in (1-3), and let $p_1, p_3 \in (1, 2]$. Then the Fourier restriction estimate (1-2) for μ is equivalent to the Fourier restriction estimate (2-3) for the measure μ_j for any $j \in \mathbb{Z}$ (as defined in (2-2)) if and only if*

$$\vartheta = \frac{|\alpha|}{p_1'} + \frac{\rho}{p_3'} - \frac{|\alpha|}{2} \tag{2-4}$$

is satisfied.

Proof. Let us first determine what 2ϑ, the degree of homogeneity of $\mathcal{W}$, needs to be in order for (2-3) to hold true for all $j \in \mathbb{Z}$ whenever it holds true for some $j_0 \in \mathbb{Z}$. Recall that $|\delta_t x|_\alpha = t|x|_\alpha$. Inspecting the definition (2-2) of μ_j one gets

$$\langle \mu_j, f \rangle = 2^{j|\alpha|+2j\vartheta} \langle \mu_0, \mathrm{Dil}_{(2^{-j\alpha_1}, 2^{-j\alpha_2}, 2^{-j\rho})} f \rangle,$$

where $(\mathrm{Dil}_{(\lambda_1,\lambda_2,\lambda_3)} f)(x_1, x_2, x_3) = f(\lambda_1^{-1}x_1, \lambda_2^{-1}x_2, \lambda_3^{-1}x_3)$. Let us assume that we have for some $j \in \mathbb{Z}$ the estimate

$$\langle \mu_j, |\hat{f}|^2 \rangle = \|\hat{f}\|^2_{L^2(\mathrm{d}\mu_j)} \le C^2 \|f\|^2_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})}.$$

Since the Fourier transform behaves well with respect to dilations $\mathrm{Dil}_{(\lambda_1,\lambda_2,\lambda_3)}$, we may rescale the above estimate and get

$$\|\hat{f}\|_{L^2(\mathrm{d}\mu_0)} \le C 2^{-j|\alpha|/2-j\vartheta+j(\alpha_1/p_1'+\alpha_2/p_1'+\rho/p_3')}\|f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})}.$$

From this one sees that we need precisely (2-4) in order for the constant in (2-3) to be independent of j. If (2-4) does not hold, then the constant blows up in one of the cases $j\to\infty$ or $j\to-\infty$, and in particular, the Fourier restriction estimate (1-2) for μ cannot hold (here we use that the restriction operators for μ and the μ_j are nonzero since $\mathcal{W}$ is not identically zero).

Let us now assume that we indeed have (2-4). It is obvious that the Fourier restriction estimate for μ implies the Fourier restriction estimate for μ_j for any j. Let us therefore assume that the estimate (2-3) holds true for any $j\in\mathbb{Z}$, and thus for all $j\in\mathbb{Z}$.

Before proceeding further let us denote by $(\tilde{\eta}_j)_{j\in\mathbb{Z}}$ a family of $C_c^\infty(\mathbb{R}^2\setminus\{0\})$ functions such that

$$\tilde{\eta}_j = \tilde{\eta}_0\circ\delta_{2^{-j}} \quad \text{for all } j\in\mathbb{Z},$$

and such that $\tilde{\eta}_j$ is equal to 1 on the support of η_j. One can for example take $\tilde{\eta}_j=\sum_{|k-j|\le N}\eta_k$ for some sufficiently large N. Let us furthermore denote by S_j the Fourier multiplier operator in $\mathbb{R}^3$ with multiplier $(\tilde{\eta}_j\otimes 1)(\xi_1,\xi_2,\xi_3)=\tilde{\eta}_j(\xi_1,\xi_2)$.

Now (2-3) implies

$$\|\widehat{S_j f}\|_{L^2(\mathrm{d}\mu_j)} = \|\hat{f}\|_{L^2(\mathrm{d}\mu_j)} \le C\|S_j f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})}.$$

Therefore

$$\begin{aligned}\|\hat{f}\|^2_{L^2(\mathrm{d}\mu)} &= \langle\mu,|\hat{f}|^2\rangle = \sum_{j\in\mathbb{Z}}\langle\mu_j,|\hat{f}|^2\rangle = \sum_{j\in\mathbb{Z}}\langle\mu_j,|\widehat{S_j f}|^2\rangle\\ &\le C^2\sum_{j\in\mathbb{Z}}\|S_j f\|^2_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})} = C^2\big\|\|S_j f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})}\big\|^2_{l^2_j},\end{aligned}$$

where l^2_j denotes the norm of the Hilbert space of l^2 sequences on $\mathbb{Z}$. Since both $p_1\le 2$ and $p_3\le 2$, we may use Minkowski's inequality to interchange the l^2_j norm with the $L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})$ norm, and subsequently apply Littlewood–Paley theory in the (x_1,x_2)-variable (in particular, we do not need to use mixed-norm Littlewood–Paley theory) to get

$$\begin{aligned}\big\|\|S_j f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})}\big\|^2_{l^2_j} &\le \big\|\|S_j f\|_{l^2_j}\big\|^2_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})}\\ &= \Big\|\Big(\sum_{j\in\mathbb{Z}}|S_j f|^2\Big)^{1/2}\Big\|^2_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})} \sim \|f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})}.\end{aligned}$$

□

Remark 2.2 (scaling in the case of Hessian determinant). Using the homogeneity condition of ϕ one easily obtains that the Hessian determinant is also α-mixed homogeneous of degree $2\rho-2|\alpha|$. Thus, when we take $\mathcal{W}=|\mathcal{H}_\phi|^\sigma$, $\mathcal{W}$ is homogeneous of degree $2\vartheta=2\sigma(\rho-|\alpha|)$. Recall that in this case (i.e., as in the assumptions of Theorem 1.1) we assume that

$$\frac{1}{p_1'}=\frac{1}{2}-\sigma, \quad \frac{1}{p_3'}=\sigma,$$

and so by (2-4) the equality $2\vartheta = 2\sigma(\rho - |\alpha|)$ is indeed satisfied, i.e., the desired relation between the exponents if one wants scaling invariance.

Remark 2.3 (a general sufficient condition for local integrability of $\mathcal{W}$). Since $\mathcal{W}$ is mixed homogeneous of degree 2ϑ, $\mathcal{W}|x|_\alpha^{-2\vartheta}$ is mixed homogeneous of degree 0, and in particular a bounded function. Thus $|\mathcal{W}| \lesssim |x|_\alpha^{2\vartheta}$, and so it is sufficient to check when $|x|_\alpha^{2\vartheta}$ is locally integrable in $\mathbb{R}^2$. By symmetry it is sufficient to integrate over $\{(x_1, x_2) : x_1, x_2 > 0\}$. We have

$$\begin{aligned}\int_{x_1,x_2>0,|x|\lesssim 1} |x|_\alpha^{2\vartheta}\,\mathrm{d}x &= \int_{x_1,x_2>0,|x|\lesssim 1} (x_1^{1/\alpha_1} + x_2^{1/\alpha_2})^{2\vartheta}\,\mathrm{d}x\\ &\sim \int_{y_1,y_2>0,|y|\lesssim 1} (y_1^2 + y_2^2)^{2\vartheta} y_1^{2\alpha_1-1} y_2^{2\alpha_2-1}\,\mathrm{d}y\\ &\sim \int_{0<r\lesssim 1}\int_0^{\pi/2} r^{4\vartheta+2|\alpha|-1}(\cos\theta)^{2\alpha_1-1}(\sin\theta)^{2\alpha_2-1}\,\mathrm{d}\theta\,\mathrm{d}r.\end{aligned}$$

Therefore, we must have $4\vartheta + 2|\alpha| - 1 > -1$, i.e.,

$$2\vartheta + |\alpha| > 0.$$

Note that this holds if $\rho \ge 0$, $p_1 > 1$, and ϑ is given by (2-4).

Remark 2.4. When ϕ is smooth at the origin and a nonconstant function, then $\rho = 1$, and the necessary condition obtained by a Knapp-type example associated to the principal face of $\mathcal{N}(\phi)$ in the initial coordinate system (see [Palle 2021, Proposition 2.1]) tells us that

$$\frac{|\alpha|}{p_1'} + \frac{1}{p_3'} \le \frac{|\alpha|}{2}$$

is necessary for (1-2) if $\mathcal{W} \equiv 1$ (i.e., $\vartheta = 0$). On the other hand, if we define $l_\alpha = \{(t_1, t_3) \in \mathbb{R}^2 : |\alpha| t_1 + t_3 = |\alpha|/2\}$, then the expression (2-4) for ϑ implies that

$$|\vartheta| = \sqrt{1 + |\alpha|^2}\,\operatorname{dist}\left(\left(\frac{1}{p_1'}, \frac{1}{p_3'}\right), l_\alpha\right).$$

2B. *Some further reductions.* According to Proposition 2.1, under the conditions of Theorem 1.1 or Theorem 1.2, we have to prove the Fourier restriction estimate for a measure defined by the mapping

$$f \mapsto \int_{\mathbb{R}^2} f(x, \phi(x))\,\eta(x)\,\mathcal{W}(x)\,\mathrm{d}x,$$

where $\eta \in C_c^\infty(\mathbb{R}^2 \setminus \{0\})$ is supported in a compact annulus centered at the origin. Note that in the case of the weight $\mathcal{W} = |\mathcal{H}_\phi|^\sigma$ (the case of Theorem 1.1) the degree of homogeneity $2\vartheta = 2\sigma(\rho - |\alpha|)$ satisfies the relation (2-4) by Remark 2.2.

Reductions for the amplitude η. One can easily show that in the context of the Fourier restriction problem we may make the following reductions. First, by reordering coordinates and/or changing their sign, and by splitting the amplitude η into functions with smaller support, we may restrict ourselves to amplitudes η with support contained in the half-plane $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 \gtrsim 1\}$. Then, by compactness, we may localize to small neighborhoods of points $v \neq 0$ having $v_1 \gtrsim 1$. Thus, one may assume that the support of η is contained in a small neighborhood of some generic point v satisfying $v_1 \sim 1$ and $|v| \lesssim 1$. In fact, compactness and changing signs if necessary implies that we may further assume that either $v_2 = 0$ or $v_2 \sim 1$.

Changing the affine terms of the phase. By the previous discussion it suffices to consider the measure

$$f \mapsto \int_{\mathbb{R}^2} f(x, \phi(x))\, \eta_v(x)\, \mathcal{W}(x)\, \mathrm{d}x, \tag{2-5}$$

where η_v is a smooth function supported in a small neighborhood of a point $v \neq 0$. We now recall the fact that we can freely add or remove linear and constant terms in the expression for ϕ in the context of the Fourier restriction problem. For the constant term this is obvious. For the linear terms this can be achieved by using a linear transformation of the form $(x_1, x_2, x_3) \mapsto (x_1, x_2, b_1 x_1 + b_2 x_2 + x_3)$ (for more details see [Palle 2021, Section 3.1] and note that here the situation is slightly simpler since no Jacobian factor appears). In particular, instead of considering the measure (2-5), we may consider the measure

$$f \mapsto \int_{\mathbb{R}^2} f(x, \phi_v(x - v))\, \eta_v(x)\, \mathcal{W}(x)\, \mathrm{d}x,$$

where we recall that

$$\phi_v(x) := \phi(x + v) - \phi(v) - x \cdot \nabla\phi(v).$$

The strategy for the proofs of Theorems 1.1 and 1.2 should now be clear. The above discussion reduces the problem to proving a local Fourier restriction estimate in the vicinity of a point v, and so one needs to determine the local normal form of ϕ at v, and in the case $\mathcal{W}(x) = |\mathcal{H}_\phi(x)|^\sigma$ one needs to additionally determine the order of vanishing of the Hessian determinant at v in the x_2-direction (after which the normal form of $\mathcal{W}$ will be clear by homogeneity).

3. Local normal forms

In this section we derive the local normal forms for ϕ and for the Hessian determinant $\mathcal{H}_\phi$ at a fixed point $v \neq 0$ (as a consequence we prove Proposition 1.4). The discussion in Section 2B implies that we may assume that $v_1 \sim 1$, and either $v_2 = 0$ or $v_2 \sim 1$.

The structure of this section is as follows. In Section 3A we fix the notation for this section, introduce relevant quantities, and define the coordinate systems y, z, and w (the coordinate systems z and w will not be described precisely until Section 3E though). In Subsections 3B, 3C, and 3D tables with normal forms of ϕ_v are given. It turns out that in most cases y-coordinates suffice and when we use them one obtains the normal forms easily. We deal with the case when y-coordinates do not suffice in Section 3E. In Section 3F we sketch how to calculate what is the order of vanishing of the Hessian determinant for the respective normal forms.

We assume that the (H1) condition is satisfied throughout this section. In fact, in Section 3B we shall explicitly determine the local normal form of ϕ when $t \mapsto \partial_2^2\phi(v_1, t)$ is flat at v_2. In this case it turns out that the Hessian determinant either does not vanish at v, or that it is flat at v. In all the other subsections we shall assume that $t \mapsto \partial_2^2\phi(v_1, t)$ is of finite type at v_2.

3A. ***Notation and some general considerations.*** Let us begin by introducing the notation. It will be useful to define

$$\gamma := \frac{\alpha_2}{\alpha_1} > 0,$$

and for the point $v = (v_1, v_2)$ (recall $v_1 \sim 1$) we define

$$t_0 := v_2 v_1^{-\gamma}.$$

Let us denote the ∂_2-derivatives of ϕ at $(1, t_0)$ by

$$b_j := \partial_2^j \phi(1, t_0) = g^{(j)}(t_0), \quad j \in \mathbb{N}_0,$$

where

$$g(t) := \phi(1, t).$$

We furthermore define

$$k := \inf\{j \geq 2 : b_j \neq 0\}, \tag{3-1}$$

where we take $k = \infty$ if $b_j = 0$ for all $j \geq 2$. The equality $k = \infty$ is equivalent to $g^{(2)}$ being flat at 0. What precisely happens when $g^{(2)}$ is flat at 0 shall be explained in Section 3B, and in the rest of the section (including this subsection) we assume that $k < \infty$, unless explicitly stated otherwise.

General form of mixed homogeneous ϕ. Recall that we denote by $\rho \in \{-1, 0, 1\}$ the degree of homogeneity of ϕ. Then we have for any x satisfying $x_1 > 0$:

$$\phi(x_1, x_2) = x_1^{\rho/\alpha_1} \phi(1, x_2 x_1^{-\gamma}). \tag{3-2}$$

Let us consider the Taylor expansion of $t \mapsto \phi(1, t)$ at t_0:

$$g(t) = \phi(1, t) = b_0 + (t - t_0) b_1 + \frac{1}{k!}(t - t_0)^k g_k(t),$$

where g_k is a smooth function such that $b_k = g_k(0)$. Thus, we get

$$\begin{aligned} \phi(x) &= x_1^{\rho/\alpha_1}\Big(b_0 + (x_2 x_1^{-\gamma} - t_0) b_1 + \frac{1}{k!}(x_2 x_1^{-\gamma} - t_0)^k g_k(x_2 x_1^{-\gamma})\Big) \\ &= x_1^{\rho/\alpha_1}(b_0 - t_0 b_1) + x_2 x_1^{(\rho-\alpha_2)/\alpha_1} b_1 + \frac{1}{k!} x_1^{(\rho-k\alpha_2)/\alpha_1}(x_2 - t_0 x_1^{\gamma})^k g_k(x_2 x_1^{-\gamma}). \end{aligned} \tag{3-3}$$

More generally, we have the formal series expansion:

$$\begin{aligned} \phi(x) &\approx \sum_{j=0}^{\infty} \frac{b_j}{j!}(x_2 - t_0 x_1^{\gamma})^j \, x_1^{\rho/\alpha_1 - j\gamma} \\ &= b_0 x_1^{\rho/\alpha_1} + b_1(x_2 - t_0 x_1^{\gamma}) x_1^{\rho/\alpha_1 - \gamma} + \sum_{j=k}^{\infty} \frac{b_j}{j!}(x_2 - t_0 x_1^{\gamma})^j \, x_1^{\rho/\alpha_1 - j\gamma}. \end{aligned} \tag{3-4}$$

If $\gamma = 1$ (i.e., $\alpha_1 = \alpha_2$) it will usually be better to write

$$\phi(x) = x_1^{\rho/\alpha_1} b_0 + (x_2 - t_0 x_1) x_1^{\rho/\alpha_1 - 1} b_1 + \frac{1}{k!}(x_2 - t_0 x_1)^k \, x_1^{\rho/\alpha_1 - k} \, g_k(x_2 \, x_1^{-1}). \tag{3-5}$$

Since $v_1 \sim 1$, we may assume

$$|x_1^{1/\alpha_1} - v_1^{1/\alpha_1}| \ll 1, \quad |x_2 \, x_1^{-\gamma} - v_2 \, v_1^{-\gamma}| \ll 1.$$

The second condition is equivalent to $|x_2 - t_0 x_1^\gamma| \ll 1$. Note that the points on the homogeneity curve through v satisfy the equation $x_2 = t_0 x_1^\gamma$.

In order to determine the normal forms it will suffice to introduce three additional coordinate systems, which we shall denote by y, z, and w respectively, each having the point v as their origin. The original coordinate system is denoted by x. The function ϕ in the coordinate system y (resp. z, w) shall be denoted by ϕ^y (resp. ϕ^z, ϕ^w). For the original coordinate system x we simply use ϕ, or ϕ^x for emphasis.

The function ϕ in the coordinate system y (resp. z, w) but without the affine terms at v shall be denoted by ϕ_v^y (resp. ϕ_v^z, ϕ_v^w). This means

$$\phi_v^y(y) := \phi^y(y) - \phi^y(0) - y \cdot \nabla\phi^y(0),$$

and similarly for ϕ_v^z and ϕ_v^w.

The coordinate system y. It is defined through the following affine coordinate change having $v = (v_1, v_2)$ as the origin:

$$\begin{aligned} y_1 &= x_1 - v_1, \\ y_2 &= x_2 - v_2 - \gamma v_2 v_1^{-1}(x_1 - v_1) \\ &= x_2 - (1-\gamma)v_2 - \gamma v_2 v_1^{-1} x_1. \end{aligned}$$

The reverse transformation is

$$\begin{aligned} x_1 &= y_1 + v_1, \\ x_2 &= y_2 + v_2 + \gamma v_2 v_1^{-1} y_1. \end{aligned} \tag{3-6}$$

One can easily check that in these coordinates we can write

$$\begin{aligned} x_2 - t_0 x_1^\gamma &= y_2 + v_2 + \gamma v_2 v_1^{-1} y_1 - v_2(1 + v_1^{-1} y_1)^\gamma \\ &= y_2 + v_2 + \gamma v_2 v_1^{-1} y_1 - v_2\Big(1 + \gamma v_1^{-1} y_1 + \binom{\gamma}{2} v_1^{-2} y_1^2 + \mathcal{O}(y_1^3)\Big) \\ &= y_2 - y_1^2 \omega(y_1); \end{aligned}$$

i.e., the points on the homogeneity curve through v satisfy the equation $y_2 = y_1^2 \omega(y_1)$ in y-coordinates. Above (and in the following) we use the notation

$$\binom{c}{m} = c(c-1) \cdots \frac{c-m+1}{m!}$$

for $c \in \mathbb{R}$ and m a nonnegative integer. Furthermore, we obviously have:

Remark 3.1. It holds that $\omega(0) \neq 0$ if and only if ω is not identically 0 if and only if $v_2 \neq 0$ (i.e., $t_0 \neq 0$) and $\gamma \neq 1$.

The coordinate system y will be used in most of the normal forms below which shall follow directly from the expression

$$\begin{aligned}\phi^y(y) = (v_1+y_1)^{\rho/\alpha_1}(b_0-t_0b_1)\\ +(v_2+y_2+\gamma v_2v_1^{-1}y_1)\,(v_1+y_1)^{(\rho-\alpha_2)/\alpha_1}b_1+(y_2-y_1^2\omega(y_1))^k r(y),\end{aligned} \tag{3-7}$$

which one obtains from (3-3) and (3-6). When $\gamma=1$, one uses (3-5) instead and gets

$$\phi^y(y) = (v_1+y_1)^{\rho/\alpha_1}b_0+y_2(v_1+y_1)^{\rho/\alpha_1-1}b_1+y_2^k r(y). \tag{3-8}$$

In both (3-7) and (3-8) the function r is smooth and nonvanishing at the origin. Let us also note that the expansion (3-4) can be rewritten in y coordinates as

$$\begin{aligned}\phi^y(y) \approx b_0(v_1+y_1)^{\rho/\alpha_1}\\ +b_1(y_2-y_1^2\omega(y_1))(v_1+y_1)^{\rho/\alpha_1-\gamma}+\sum_{j=k}^{\infty}\frac{b_j}{j!}(y_2-y_1^2\omega(y_1))^j\,(v_1+y_1)^{\rho/\alpha_1-j\gamma}.\end{aligned} \tag{3-9}$$

The following simple lemma shall be useful later:

Lemma 3.2. *From (3-7) and (3-8) we get the following information on the second-order derivatives of ϕ^y:*

(1) *It always holds*

$$k=2 \iff b_2\neq 0 \iff \partial_2^2\phi^y(0)\neq 0.$$

(2.a) *If $\rho\neq 1$ or $\alpha_2\neq 1$ (i.e., $\rho-\alpha_2\neq 0$), then*

$$b_1\neq 0 \iff \partial_1\partial_2\phi^y(0)\neq 0.$$

(2.b) *If $\rho=\alpha_2=1$ or if $b_1=0$, then $\partial_1\partial_2\phi^y(0)=0$.*

(3.a) *If $\rho=0$ and $\alpha_1\neq\alpha_2$ (i.e., $\gamma\neq 1$), or if $\rho=\alpha_1=1$ and $\alpha_2\neq 1$ (and in particular $\gamma\neq 1$), then*

$$b_1\neq 0,\ \ t_0\neq 0 \iff \partial_1^2\phi^y(0)\neq 0,$$

and recall that $v_2\neq 0$ if and only if $t_0\neq 0$.

(3.b) *If $\rho=\alpha_2=1$ and $\alpha_1\neq 1$ (and in particular $\gamma\neq 1$), then*

$$b_0-t_0b_1\neq 0 \iff \partial_1^2\phi^y(0)\neq 0.$$

(3.c) *If $\gamma=1$ (i.e., $\alpha_1=\alpha_2$) or if $b_1=0$, then*

$$b_0\neq 0,\ \frac{\rho}{\alpha_1}\notin\{0,1\} \iff \partial_1^2\phi^y(0)\neq 0.$$

Note that $\rho/\alpha_1=0$ if and only if $\rho=0$, and $\rho/\alpha_1=1$ if and only if $\rho=\alpha_1=1$.

Proof. The only not completely trivial case is (3.a). Since in this case $\rho/\alpha_1\in\{0,1\}$, the first term in (3-7) is an affine term, and so we can ignore it. Since $k\geq 2$, the third term also does not contribute to the

y_1^2-term in the Taylor series of ϕ^y, and so we can ignore it too. We therefore only need to consider the term

$$(v_2 + y_2 + \gamma v_2 v_1^{-1} y_1)\,(v_1 + y_1)^{(\rho-\alpha_2)/\alpha_1} b_1,$$

and in fact, we may even reduce ourselves to

$$(v_2 + \gamma v_2 v_1^{-1} y_1)\,(v_1 + y_1)^{(\rho-\alpha_2)/\alpha_1} b_1 = b_1 v_2 (1 + \gamma v_1^{-1} y_1)\,(v_1 + y_1)^{(\rho-\alpha_2)/\alpha_1}.$$

Now if $t_0 = 0$ (i.e., $v_2 = 0$) or if $b_1 = 0$, then $\partial_1^2 \phi^y(0) = 0$ follows. Let us now assume $v_2 \neq 0$ and $b_1 \neq 0$. We note that in our case we may rewrite $(\rho - \alpha_2)/\alpha_1 = \rho - \gamma$, and so it suffices to show that

$$\partial_{y_1}^2|_{y_1=0}((1 + \gamma v_1^{-1} y_1)\,(1 + v_1^{-1} y_1)^{\rho-\gamma}) \neq 0.$$

Calculating the second derivative one gets

$$2\gamma v_1^{-2}(\rho - \gamma) + v_1^{-2}(\rho - \gamma)(\rho - \gamma - 1).$$

This is not zero since in this case we have $\rho \in \{0, 1\}$ and $\gamma \notin \{0, 1\}$. $\square$

The coordinate systems z and w. These are defined through affine coordinate changes of the form

$$\begin{aligned} x_1 &= v_1 + z_1, & w_1 &= z_1 + \frac{1}{B} z_2, \\ x_2 &= v_2 + z_2 + A z_1, & w_2 &= z_2, \end{aligned} \tag{3-10}$$

having (v_1, v_2) as their origin, where we shall have $B := A - \gamma v_2 v_1^{-1} \neq 0$ so that the coordinate system y never coincides with the coordinate system z, and the coordinate system z never coincides with the coordinate system w. The constant A shall depend on v and the first few derivatives of ϕ at v (note that $A = B \neq 0$ if $v_2 = t_0 = 0$). These coordinate systems will be described more precisely in Section 3E. There we shall also introduce a smooth function $\tilde{\omega}$ such that

$$x_2 - t_0 x_1^\gamma = y_2 - y_1^2 \omega(y_1) = (w_1 - w_2^2 \tilde{\omega}(w_2)) r_0(w)$$

for some smooth function r_0 satisfying $r_0(0) \neq 0$. Note that we have

$$\begin{aligned} y_1 &= z_1 = w_1 - \frac{1}{B} w_2, \\ y_2 &= z_2 + B z_1 = B w_1. \end{aligned} \tag{3-11}$$

As we shall see in Section 3E below the z-coordinates are only used in the intermediate steps and the normal forms are expressed exclusively in y- or w-coordinates.

Some general considerations regarding the Hessian determinant $\mathcal{H}_\phi$. Recall that

$$\phi(t^{\alpha_1} x_1, t^{\alpha_2} x_2) = t^\rho \phi(x_1, x_2).$$

Taking derivatives in x_1 and x_2 we get

$$(\partial_1^{\tau_1} \partial_2^{\tau_2} \phi)(t^{\alpha_1} x_1, t^{\alpha_2} x_2) = t^{\rho - \tau_1\alpha_1 - \tau_2\alpha_2} (\partial_1^{\tau_1} \partial_2^{\tau_2} \phi)(x_1, x_2).$$

Thus, we have for the Hessian determinant of ϕ:

$$\mathcal{H}_\phi(t^{\alpha_1}x_1, t^{\alpha_2}x_2) = t^{2(\rho-|\alpha|)}\mathcal{H}_\phi(x_1, x_2).$$

From this it follows that if $\mathcal{H}_\phi$ vanishes at the point v, then it also vanishes along the homogeneity curve through v, which we recall is parametrized by $t \mapsto (t^{\alpha_1}v_1, t^{\alpha_2}v_2)$.

We are interested in the order of vanishing of $\mathcal{H}_\phi$ in directions transversal to this curve. In particular, if we have $\partial_2^{\tau_2}\mathcal{H}_\phi(v) = 0$ for $\tau_2 < N$ and $\partial_2^N\mathcal{H}_\phi(v) \neq 0$, then by using homogeneity and a Taylor expansion (as we did for ϕ) we get

$$\mathcal{H}_\phi(x) = (x_2 - t_0x_1^\gamma)^N r_0(x)$$

for some smooth function r_0 satisfying $r_0(v) \neq 0$. Calculating N shall be done in Section 3F by using the normal forms of ϕ. Recall that the Hessian determinant is equivariant under affine coordinate changes, and so we can freely change to y-, z-, or w-coordinates.

Preliminary comments on the normal forms. Let us introduce the following notation for the nondegenerate case (i.e., the case when the Hessian determinant of ϕ does not vanish at v):

(ND) The function ϕ_v is nondegenerate at the origin.

When ϕ_v does not satisfy (ND), we note that Proposition 1.4 implies in particular that after a linear change of coordinates the function ϕ_v takes one of the following three forms:

$$\phi_v^u(u) = u_1^{k_0}r(u) + \varphi(u),$$
$$\phi_v^u(u) = u_1^2r_1(u) + u_2^{k_0}r_2(u),$$
$$\phi_v^u(u) = (u_2 - u_1^2\psi(u_1))^{k_0}r(u),$$

where $r(0), \psi(0), r_1(0), r_2(0) \neq 0$, φ is flat at 0, and $k_0 \geq 2$ in the first and third cases, while $k_0 \geq 3$ in the second. Note that the first case corresponds to the normal form (i) of Proposition 1.4, the second case is a reduced version of normal forms (ii), (iii), (iv), (v), and the third corresponds to the normal form (vi).

However, the above three forms do not contain sufficient information to obtain restriction estimates. In this section we shall obtain the much more detailed classification given in Table 2.

All the appearing functions are smooth and do not vanish at the origin, except the function φ, which is always flat at the origin. The number k is as defined in (3-1) and it is always finite in the above normal forms (when it is infinite it turns out that one is necessarily in the case of normal form (i.y2)). On the other hand, the definition of the number $\tilde{k}$ changes from case to case, and we allow $\tilde{k}$ to be infinite only in normal form (i.y1), in which case we consider the Hessian determinant to be flat at the origin. The quantities v_1, γ, A, B appearing in the conditions column and the functions ω and $\tilde{\omega}$ are the same ones as previously defined in this subsection. Let us furthermore remark that normal forms (i.w1) and (i.w2) stem from normal forms (ii.w), (iii), and (v), in the sense that they correspond to $\tilde{k} = \infty$.

Two remarks before we continue. First, note that the normal forms listed in Proposition 1.4 are a compressed version of Table 2 — in the proposition we ignored the subcases, e.g., the normal forms (i.y1), (i.y2), (i.w1), (i.w2) were all compressed in Proposition 1.4 to a single normal form (i). Second,

case	normal form	additional conditions
(i.y1)	$\phi_v^y(y) = y_2^k r(y)$, $\mathcal{H}_{\phi^y}(y) = y_2^{\tilde{k}+2k-2} r_0(y)$	$k \geq 2$, $\tilde{k} \geq 0$ or $\tilde{k} = \infty$
(i.y2)	$\phi_v^y(y) = y_1^{\tilde{k}} q(y_1) + \varphi(y)$, $\mathcal{H}_{\phi^y}$ is flat at 0	$\tilde{k} \geq 2$
(i.w1)	$\phi_v^w(w) = w_2^2 q(w_2) + \varphi(w)$, $\mathcal{H}_{\phi^w}$ is flat at 0	-
(i.w2)	$\phi_v^w(w) = w_2^2 r(w) + \varphi(w)$, $\mathcal{H}_{\phi^w}$ is flat at 0	$v_1 B \partial_1^j r(0) = jA(\gamma - 1)\partial_1^{j-1} r(0)$ for all $j \geq 1$
(ii.y)	$\phi_v^y(y) = y_1^2 q(y_1) + y_2^k r(y)$, $\mathcal{H}_{\phi^y}(y) = y_2^{k-2} r_0(y)$	$k \geq 3$
(ii.w)	$\phi_v^w(w) = w_1^{\tilde{k}} r(w) + w_2^2 q(w_2)$, $\mathcal{H}_{\phi^w}(w) = w_1^{\tilde{k}-2} r_0(w)$	$\tilde{k} \geq 3$
(iii)	$\phi_v^w(w) = w_1^{\tilde{k}} r_1(w) + w_2^2 r_2(w)$, $\mathcal{H}_{\phi^w}(w) = w_1^{\tilde{k}-2} r_0(w)$	$\tilde{k} \geq 3$, $v_1 B \partial_1^j r_2(0) = jA(\gamma - 1)\partial_1^{j-1} r_2(0)$ for all $1 \leq j \leq \tilde{k} - 1$
(iv)	$\phi_v^y(y) = y_1^2 q(y_1) + (y_2 - y_1^2 \omega(y_1))^k r(y)$, $\mathcal{H}_{\phi^y}(y) = (y_2 - y_1^2 \omega(y_1))^{k-2} r_0(y)$	$k \geq 3$
(v)	$\phi_v^w(w) = (w_1 - w_2^2 \tilde{\omega}(w_2))^{\tilde{k}} r_1(w) + w_2^2 r_2(w)$, $\mathcal{H}_{\phi^w}(w) = (w_1 - w_2^2 \tilde{\omega}(w_2))^{\tilde{k}-2} r_0(w)$	$\tilde{k} \geq 3$, $v_1 B \partial_1^j r_2(0) = jA(\gamma - 1)\partial_1^{j-1} r_2(0)$ for $1 \leq j \leq \tilde{k} - 1$
(vi)	$\phi_v^y(y) = (y_2 - y_1^2 \omega(y_1))^k r(y)$, $\mathcal{H}_{\phi^y}(y) = (y_2 - y_1^2 \omega(y_1))^{2k-3} r_0(y)$	$k \geq 2$

Table 2. Detailed classification of normal forms (an uncompressed version of Table 1).

note that the above “uncompressed” table of normal forms is not mutually exclusive in the sense that the forms themselves differ from each other — for example in this sense the normal form (i.y2) obviously contains the case of the normal form (i.w1), the main difference being only the coordinate system which one needs to use in order to obtain them. On the other hand, the normal forms in the compressed table in Proposition 1.4 are in this sense indeed mutually exclusive.

The first step in deriving the above normal forms is to switch to y-coordinates. In most cases (see the tables of cases below for the precise list) this will suffice and the normal form will be obvious, and so in the following subsections we shall leave out most of the details for them. In particular, as a consequence of considerations in Subsections 3C and 3D, we shall obtain:

Lemma 3.3. *If $k \geq 3$ and if we are not in the* (ND) *case, then the function ϕ_v^y is always in one of the normal forms* (i.y1), (i.y2), (ii.y), (iv), *or* (vi).

If $k = 2$, $b_1 \neq 0$, $\rho \neq \alpha_2$, and we are not in the (ND) case, then we shall either need to

(FP) flip coordinates (i.e., exchange x_1 and x_2) and use the y-coordinates associated to the flipped coordinates,

or we shall need w- (and the intermediary z-) coordinates. Details can be found in Section 3E below.

Note that flipping coordinates makes sense only when $v_2 \neq 0$ (and indeed, we shall flip coordinates only when $A = 0$, which, as it turns out, never happens when $v_2 = 0$). After flipping coordinates it will always suffice to use the y-coordinates (associated to the flipped x, v, and α), and in particular, we shall be able to apply Lemma 3.3. Note that these y coordinates are not in general equal to flipped y-coordinates associated to the original x, v, and α.

3B. *Normal form when $t \mapsto \partial_2^2\phi(1,t)$ is flat at t_0 (i.e., $k = \infty$).* Let us assume that

$$\partial_2^j \phi(1, t_0) = 0 \quad \text{for all } j \geq 2, \tag{3-12}$$

and so we have $\partial_2^j \phi(v) = 0$ for all v (with $v_1 > 0$) satisfying $v_2 v_1^{-\gamma} = t_0$ by (3-2). By the Euler homogeneous function theorem ϕ satisfies the equation

$$\rho\phi(x) = \alpha_1 x_1 \partial_1 \phi(x) + \alpha_2 x_2 \partial_2 \phi(x).$$

Taking the derivative $\partial^\tau = \partial_1^{\tau_1} \partial_2^{\tau_2}$ we get at (v_1, v_2) that

$$(\rho - \alpha_1\tau_1 - \alpha_2\tau_2)\partial^\tau \phi(v) = \alpha_1 v_1 \partial^{\tau+(1,0)} \phi(v) + \alpha_2 v_2 \partial^{\tau+(0,1)} \phi(v).$$

From this, the fact that $\alpha_1 v_1 \neq 0$, and the flatness assumption (3-12) it follows by induction in τ_1 that $\partial^\tau \phi(v) = 0$ for all $\tau_1 \geq 0$ and $\tau_2 \geq 2$.

If now $\partial_1 \partial_2 \phi(v) \neq 0$, then the Hessian determinant does not vanish and we are in the (ND) case (this always happens for example when $\phi(x_1, x_2) = x_1 x_2$). On the other hand, if $\partial_1 \partial_2 \phi(v) = 0$, then we get in the same way as above that $\partial^\tau \phi(v) = 0$ for all $\tau_1 \geq 1$ and $\tau_2 = 1$. Thus, by using a Taylor expansion at v and by switching to y-coordinates (recall $x_1 = y_1 + v_1$) we may write

$$\phi_v^y(y) = y_1^2 q(y_1) + \varphi(y),$$

where q is a smooth function and φ is a smooth function flat at the origin. In particular, in this case the Hessian determinant vanishes of infinite order at $x = v$ and therefore the condition (H2) cannot hold. This also shows that (H2) is a stronger condition than (H1). Since we assume that at least (H1) holds, we necessarily have that $t \mapsto \partial_1^2\phi(t, v_2)$ is not flat at v_1, and so q cannot be flat at the origin either, i.e., we can write

$$\phi_v^y(y) = y_1^{\tilde{k}}\tilde{q}(y_1) + \varphi(y)$$

for some smooth function $\tilde{q}$ satisfying $\tilde{q}(0) \neq 0$ and $\tilde{k} \geq 2$. This is precisely the normal form (i.y2).

3C. ***Normal form tables for ϕ mixed homogeneous of degree $\rho = 0$.*** Recall that we assume $k < \infty$ in this and the following subsections. In this case (3-7) becomes

$$\phi^y(y) - (b_0 - t_0 b_1) = (v_2 + y_2 + \gamma v_2 v_1^{-1} y_1)\,(v_1 + y_1)^{-\gamma} b_1 + (y_2 - y_1^2\omega(y_1))^k r(y)$$

if $\gamma \neq 1$, and in the case $\gamma = 1$ we have by (3-8) that

$$\phi^y(y) - b_0 = y_2(v_1 + y_1)^{-1} b_1 + y_2^k r(y). \tag{3-13}$$

We have put the constant terms on the left-hand side since we may freely ignore them. Note that in the case $\gamma = 1$ we have $\partial_1^2\phi^y(0) = 0$ by Lemma 3.2 (3.c).

Case $\gamma = 1$.

conditions	case
$b_1 = 0$	normal form (i.y1)
$b_1 \neq 0$	(ND)

Here we actually have in the case when $b_1 = 0$ a precise order of vanishing of the Hessian determinant: it is always $2k - 2$. This follows from Section 3F (see in particular (3-31)).

If $b_1 \neq 0$, then from (3-13) we obviously have $\partial_1\partial_2\phi^y(0) \neq 0$, and it follows that the Hessian determinant at 0 is nonzero.

Case $\gamma \neq 1$.

conditions	case
$t_0 = 0,\ b_1 = 0$	normal form (i.y1)
$t_0 = 0,\ b_1 \neq 0$	(ND)
$t_0 \neq 0,\ b_1 = 0$	normal form (vi)
$t_0 \neq 0,\ b_1 \neq 0,\ k \geq 3$	(ND)
$t_0 \neq 0,\ b_1 \neq 0,\ k = 2$	(ND), or (FP), or normal form (v), or normal form (i.w2)

In the case $t_0 = 0,\ b_1 \neq 0$ we apply Lemma 3.2 (2.a) and (3.a), and get respectively that $\partial_1\partial_2\phi^y(0) \neq 0$ and $\partial_1^2\phi^y(0) = 0$, from which it indeed follows that we are in the (ND) case. Similarly, in the case $t_0 \neq 0$, $b_1 \neq 0,\ k \geq 3$ we use Lemma 3.2 (1) and (2.a), and obtain that $\partial_2^2\phi^y(0) = 0$ and $\partial_1\partial_2\phi^y(0) \neq 0$, from which we again get that the Hessian determinant of ϕ^y does not vanish.

As the case $t_0 \neq 0,\ b_1 \neq 0,\ k = 2$ shall be treated in the same way as certain other cases which appear later and where w-coordinates may be needed, we have postponed its discussion to Section 3E.

3D. ***Normal form tables for ϕ mixed homogeneous of degree $\rho = \pm 1$.*** Recall that by (3-3) here we have

$$\phi(x) = x_1^{\rho/\alpha_1}(b_0 - t_0 b_1) + x_2\, x_1^{(\rho-\alpha_2)/\alpha_1} b_1 + \frac{1}{k!} x_1^{(\rho-k\alpha_2)/\alpha_1}(x_2 - t_0 x_1^{\gamma})^k g_k(x_2\, x_1^{-\gamma})$$

and according to (3-7) in y-coordinates this is

$$\phi^y(y) = (v_1+y_1)^{\rho/\alpha_1}(b_0 - t_0 b_1) + (v_2 + y_2 + \gamma v_2 v_1^{-1} y_1)\,(v_1+y_1)^{(\rho-\alpha_2)/\alpha_1} b_1 + (y_2 - y_1^2\omega(y_1))^k r(y).$$

In this subsection (where $\rho = \pm 1$) we need to consider five possible subcases. The cases we first consider are when $\rho = \alpha_1$, or $\rho = \alpha_2$, or both. Since α_1 and α_2 are strictly positive, these cases are only possible for $\rho = 1$. The penultimate case is when $\alpha_1 = \alpha_2 \neq \rho$, and the last case is when all of α_1, α_2, and ρ are different from each other.

Case $\rho = 1$, $\alpha_1 = 1$, $\alpha_2 = 1$. In this case the first two terms in (3-7) become affine, and by Remark 3.1 we have $\omega \equiv 0$. As a consequence we have only one case.

conditions	case
-	normal form (i.y1)

Furthermore, we note that initially we know that the order of vanishing of the Hessian determinant is at least $2k-2$, which is always greater than or equal to 2. Since this is true at every point, the Hessian determinant vanishes identically in this case.

Case $\rho = 1$, $\alpha_1 \neq 1$, $\alpha_2 = 1$. Here we first note that by Lemma 3.2 (2.b), we always have $\partial_1\partial_2\phi^y(0) = 0$. This is a simple consequence of the fact that in this case the second term in (3-7) is linear.

conditions	case
$b_0 - t_0 b_1 = 0,\ t_0 = 0$	normal form (i.y1)
$b_0 - t_0 b_1 = 0,\ t_0 \neq 0$	normal form (vi)
$b_0 - t_0 b_1 \neq 0,\ k = 2$	(ND)
$b_0 - t_0 b_1 \neq 0,\ k \geq 3,\ t_0 = 0$	normal form (ii.y)
$b_0 - t_0 b_1 \neq 0,\ k \geq 3,\ t_0 \neq 0$	normal form (iv)

The first two cases in the table are now clear since the first two terms in (3-7) can be ignored. The (ND) case follows from Lemma 3.2 (1) and (3.b) (and as previously mentioned (2.b)). The last two cases follow simply by developing the first term in (3-7) in a Taylor series in y_1 and ignoring the constant and the linear term.

Case $\rho = 1$, $\alpha_1 = 1$, $\alpha_2 \neq 1$. Here we note that the first term in (3-7) becomes linear, and therefore does not influence the normal form of ϕ_v^y.

conditions	case
$t_0 = 0,\ b_1 = 0$	normal form (i.y1)
$t_0 = 0,\ b_1 \neq 0$	(ND)
$t_0 \neq 0,\ b_1 = 0$	normal form (vi)
$t_0 \neq 0,\ b_1 \neq 0,\ k \geq 3$	(ND)
$t_0 \neq 0,\ b_1 \neq 0,\ k = 2$	(ND) or (FP)

The cases $t_0 = 0,\ b_1 \neq 0$ and $t_0 \neq 0,\ b_1 \neq 0, k \geq 3$ are (ND) by the same argumentation as in the table above for $\rho = 0,\ \gamma \neq 1$ (namely, by applying Lemma 3.2 (2.a) and (3.a), in the case $t_0 = 0,\ b_1 \neq 0$, and by applying Lemma 3.2 (1) and (2.a), in the case $t_0 \neq 0,\ b_1 \neq 0,\ k \geq 3$).

Let us note the following for the last case where $t_0 \neq 0,\ b_1 \neq 0$, and $k = 2$. The expression in (3-3) can be rewritten as (after ignoring the first term, which is linear in this case):

$$b_1 x_2 x_1^{1-\gamma} + \tfrac{1}{2} b_2 x_1^{1-2\gamma} (x_2 - t_0 x_1^{\gamma})^2 + \mathcal{O}((x_2 - t_0 x_1^{\gamma})^3).$$

We want to calculate what the Hessian determinant of $\phi_v^x = \phi_v$ at v is (or equivalently, the Hessian determinant of ϕ at v). For this we only need the second derivatives of ϕ at v, and so we can freely ignore the last term of size $(x_2 - t_0 x_1^{\gamma})^3$. After expanding the second term in the above expression and ignoring the linear terms and the term $\mathcal{O}((x_2 - t_0 x_1^{\gamma})^3)$ we get

$$(b_1 - t_0 b_2) x_1^{1-\gamma} x_2 + \tfrac{1}{2} b_2 x_1^{1-2\gamma} x_2^2.$$

From this it follows by a direct calculation that

$$\partial_1^2 \phi(v) = -\gamma \frac{v_2}{v_1} \partial_1 \partial_2 \phi(v),$$

and so

$$\mathcal{H}_\phi(v) = -\partial_1 \partial_2 \phi(v) \Big(\partial_1 \partial_2 \phi(v) + \gamma \frac{v_2}{v_1} \partial_2^2 \phi(v) \Big),$$

which we note can be rewritten as

$$\mathcal{H}_\phi(v) = -\partial_1 \partial_2 \phi^x(v)\, \partial_1 \partial_2 \phi^y(0),$$

by (3-6). This implies in particular that $\mathcal{H}_\phi(v) = 0$ if and only if $\partial_1 \partial_2 \phi(v) = 0$ if and only if $\partial_1^2 \phi(v) = 0$ since by Lemma 3.2 (2.a), we know that $\partial_1 \partial_2 \phi^y(0) \neq 0$.

Thus, in the last case where $t_0 \neq 0$, $b_1 \neq 0$, and $k = 2$, we are either in the (ND) case, and otherwise we have $\partial_1^2 \phi(v) = 0$. This means precisely that the "k" associated to the flipped coordinates (and we can flip coordinates since $t_0 \neq 0$, i.e., $v_2 \neq 0$) is necessarily ≥ 3. For the flipped coordinates we may now use the previous table where we have $\rho = 1,\ \alpha_1 \neq 1,\ \alpha_2 = 1$ (or apply Lemma 3.3).

Case $\rho = \pm 1,\ \alpha_1 = \alpha_2 \neq \rho$. Here one uses (3-8):

$$\phi^y(y) = (v_1 + y_1)^{\rho/\alpha_1} b_0 + y_2 (v_1 + y_1)^{\rho/\alpha_1 - 1} b_1 + y_2^k r(y).$$

conditions	case
$b_0 = 0,\ b_1 = 0$	normal form (i.y1)
$b_0 = 0,\ b_1 \neq 0$	(ND)
$b_0 \neq 0,\ b_1 = 0,\ k \geq 3$	normal form (ii.y)
$b_0 \neq 0,\ b_1 = 0,\ k = 2$	(ND)
$b_0 \neq 0,\ b_1 \neq 0,\ k \geq 3$	(ND)
$b_0 \neq 0,\ b_1 \neq 0,\ k = 2$	(ND), or (FP), or normal form (ii.w), or normal form (i.w1)

The first (ND) case $b_0 = 0,\ b_1 \neq 0$ follows from Lemma 3.2 (2.a) and (3.c), the second (ND) case $b_0 \neq 0,\ b_1 = 0,\ k = 2$ follows from Lemma 3.2 (2.a), (3.c), and (1), and the third (ND) case $b_0 \neq 0$, $b_1 \neq 0,\ k \geq 3$ follows from Lemma 3.2 (1) and (2.a). For the last case $b_0 \neq 0,\ b_1 \neq 0,\ k = 2$ we again refer the reader to Section 3E.

We give two further remarks. Firstly, one can show that in the case $b_0 = 0,\ b_1 = 0$ the order of vanishing of the Hessian determinant is precisely equal to $2k-2$ if and only if we additionally have

$$\frac{\rho}{\alpha_1} \notin \{1, k\},$$

as is shown in Section 3F. Note that here we cannot have $\rho/\alpha_1 = 1$, and when $\rho/\alpha_1 = k$ from Section 3F we see that the Hessian determinant vanishes of order $2k + \tilde{k} - 2$, where $\tilde{k}$ is the smallest positive integer such that $b_{k+\tilde{k}} \neq 0$ (it is also possible $\tilde{k} = \infty$ with the obvious interpretation).

Secondly, here we can calculate explicitly from the derivatives $b_{\tau_2} = g^{(\tau_2)}(t_0)$ the number $\tilde{k}$ in the normal form (ii.w) (see (3-26) in Section 3E). This is already known for homogeneous polynomials [Ferreyra et al. 2004].

Case $\boldsymbol{\rho = \pm 1,\ \alpha_1 \neq \rho,\ \alpha_2 \neq \rho,\ \alpha_1 \neq \alpha_2}$.

conditions	case
$b_1 = 0,\ b_0 = 0,\ t_0 = 0$	normal form (i.y1)
$b_1 = 0,\ b_0 = 0,\ t_0 \neq 0$	normal form (vi)
$b_1 = 0,\ b_0 \neq 0,\ k = 2$	(ND)
$b_1 = 0,\ b_0 \neq 0,\ k \geq 3,\ t_0 = 0$	normal form (ii.y)
$b_1 = 0,\ b_0 \neq 0,\ k \geq 3,\ t_0 \neq 0$	normal form (iv)
$b_1 \neq 0,\ k \geq 3$	(ND)
$b_1 \neq 0,\ k = 2,\ t_0 = 0$	(ND), or normal form (iii), or normal form (i.w2)
$b_1 \neq 0,\ k = 2,\ t_0 \neq 0$	(ND), or (FP), or normal form (v), or normal form (i.w2)

The first (ND) case $b_1 = 0,\ b_0 \neq 0,\ k = 2$ follows from Lemma 3.2 (1), (2.a), and (3.c), and the second (ND) case $b_1 \neq 0,\ k \geq 3$ from Lemma 3.2 (1) and (2.a). For the very last two cases (namely, $b_1 \neq 0,\ k = 2,\ t_0 = 0$ and $b_1 \neq 0,\ k = 2,\ t_0 \neq 0$) we refer the reader, as usual, to Section 3E.

Note that at this point our considerations have proven Lemma 3.3, except for the Hessian part.

3E. ***The case when $\rho \neq \alpha_2$, $b_1 \neq 0$, $k = 2$.*** In this subsection we shall discuss the remaining cases where y-coordinates did not suffice and all of which (as one easily sees from the tables in the previous two subsections) satisfy $\rho \neq \alpha_2$, $b_1 \neq 0$, $k = 2$. Here it will turn out that we are either in the (ND) case, or the (FP) case, or that we need to use the w-coordinates. In this case the form of the function ϕ in y-coordinates is according to (3-7) equal to

$$\phi^y(y) = (v_1+y_1)^{\rho/\alpha_1}(b_0 - t_0 b_1) + (v_2 + y_2 + \gamma v_2 v_1^{-1} y_1)\,(v_1+y_1)^{(\rho-\alpha_2)/\alpha_1} b_1 + (y_2 - y_1^2 \omega(y_1))^2 r(y),$$

where $r(0) \neq 0$, and, as noted in Remark 3.1, $\omega \equiv 0$ if and only if $\gamma = 1$ or $t_0 = 0$, and otherwise $\omega(0) \neq 0$. By Lemma 3.2 (1) and (2.a), we have

$$\partial_2^2 \phi^y(0) \neq 0 \quad \text{and} \quad \partial_1 \partial_2 \phi^y(0) \neq 0,$$

i.e., the y_2^2-term and the $y_1 y_2$-term in Taylor expansion of ϕ^y do not vanish. Therefore, depending on what the coefficient of the y_1^2-term is, it can happen that the Hessian determinant vanishes or not.

Case (ND) and the definition of z-coordinates. If the Hessian determinant does not vanish, we are in the nondegenerate case. Otherwise, if the Hessian determinant does vanish, then since $\partial_2^2 \phi(v) \neq 0$ (which is by definition equivalent to $k = 2$), there is a coordinate system of the form

$$x_1 = v_1 + z_1,$$
$$x_2 = v_2 + z_2 + A z_1,$$

with A unique, such that $\phi^x(x) = \phi^z(z)$, and such that the z_1^2- and $z_1 z_2$-terms in Taylor expansion of ϕ^z at the origin vanish, i.e.,

$$\partial_1^2 \phi^z(0) = 0 \quad \text{and} \quad \partial_1 \partial_2 \phi^z(0) = 0.$$

In particular, the coordinate systems y and z cannot coincide since the term $y_1 y_2$ does not vanish. This implies $B := A - \gamma v_2 v_1^{-1} \neq 0$ (compare (3-6) and (3-10)).

Case (FP) and the reduction to $A \neq 0$. Let us now prove that we may reduce ourselves to the case

$$A \neq 0.$$

If $t_0 = 0$ (i.e., $v_2 = 0$), then we always have $A = B \neq 0$. The second possibility is $t_0 \neq 0$, and if in this case we would have $A = 0$, then z- and x-coordinates would coincide (up to a translation) which implies $\partial_{x_1}^2 \phi^x(v) = \partial_{z_1}^2 \phi^z(0) = 0$. Thus, by flipping coordinates, we would have that the k associated to the flipped coordinates is ≥ 3, and so we would be in the case where the y-coordinates associated to the flipped coordinates would suffice; i.e., we could apply Lemma 3.3.

This is also the reason why in the case when $\rho = 1$, $\alpha_1 = 1$, and $\alpha_2 \neq 1$, it always sufficed to flip coordinates. The calculation below the corresponding table in Section 3D shows that $\mathcal{H}_\phi(v) = 0$ implies $\partial_1^2 \phi(v) = \partial_1 \partial_2 \phi(v) = 0$, which in turn implies that one always has $A = 0$.

The normal form in z-coordinates. Now that we may assume $A \neq 0$, our first step is to write down the Euler equation for homogeneous functions in z-coordinates. The Euler equation is

$$\rho\phi(x) = \alpha_1 x_1 \partial_1 \phi(x) + \alpha_2 x_2 \partial_2 \phi(x).$$

By the definition of z-coordinates we have

$$\partial_{x_1} = \partial_{z_1} - A\partial_{z_2} \quad \text{and} \quad \partial_{x_2} = \partial_{z_2}.$$

Thus, the Euler equation in z-coordinates is

$$\begin{aligned}\rho\phi^z(z) &= \alpha_1(v_1+z_1)\partial_1\phi^z(z) - \alpha_1 A(v_1+z_1)\partial_2\phi^z(z) + \alpha_2(v_2+z_2+Az_1)\partial_2\phi^z(z)\\ &= \alpha_1(v_1+z_1)\partial_1\phi^z(z) + (-\alpha_1 v_1 B + A(-\alpha_1+\alpha_2)z_1 + \alpha_2 z_2)\partial_2\phi^z(z).\end{aligned} \tag{3-14}$$

We now claim that if $\partial_1^{\tau_1+1}\phi^z(0) = \partial_1^{\tau_1}\partial_2\phi^z(0) = 0$ for all $1 \le \tau_1 < N$ for some $N \ge 2$, then $\partial_1^{N+1}\phi^z(0) = 0$ if and only if $\partial_1^N\partial_2\phi^z(0) = 0$. But this is almost obvious. Namely, we just take the derivative ∂_1^N at 0 in the above Euler equation and get

$$\rho\partial_1^N\phi^z(0) = \alpha_1 v_1 \partial_1^{N+1}\phi^z(0) + \alpha_1 N \partial_1^N\phi^z(0) - \alpha_1 v_1 B \partial_1^N\partial_2\phi^z(0) + AN(-\alpha_1+\alpha_2)\partial_1^{N-1}\partial_2\phi^z(0).$$

Using the assumption on vanishing derivatives we get

$$\partial_1^{N+1}\phi^z(0) = B\partial_1^N\partial_2\phi^z(0). \tag{3-15}$$

As we noted above $B \neq 0$ and our claim follows.

Now recall that $\partial_1^2\phi^z(0) = 0$ and $\partial_1\partial_2\phi^z(0) = 0$. Thus, the previously proved claim implies in particular by an inductive argument in N that either there is a $\tilde{k} \in \mathbb{N}$ such that $3 \le \tilde{k} < \infty$, satisfying

$$\tilde{k} = \min\{j \ge 2 : \partial_1^j\phi^z(0) \neq 0\} = \min\{j \ge 2 : \partial_1^{j-1}\partial_2\phi^z(0) \neq 0\},$$

and

$$\phi_v^z(z) = z_1^{\tilde{k}} r_1(z) + z_1^{\tilde{k}-1} z_2 r_2(z) + z_2^2 r_3(z), \tag{3-16}$$

where $r_i(0) \neq 0,\ i = 1,2,3$, or that

$$\phi_v^z(z) = z_1^N r_{N,1}(z) + z_1^{N-1} z_2 r_{N,2}(z) + z_2^2 r_3(z)$$

for any $N \in \mathbb{N}$, which we shall consider as the case when $\tilde{k} = \infty$. Here the $r_{N,*}$ are zero at the origin.

The normal form in w-coordinates. It will be advantageous to use w-coordinates where, unlike in (3-16), the $w_1^{\tilde{k}-1}w_2$-term is no longer present; i.e., we may write

$$\phi_v^w(w) = w_1^{\tilde{k}} r_1(w) + w_2^2 r_2(w). \tag{3-17}$$

This fact follows directly from (3-15) and from

$$\partial_{w_1} = \partial_{z_1} \quad \text{and} \quad \partial_{w_2} = \partial_{z_2} - \frac{1}{B}\partial_{z_1},$$

which we get from the definition of w coordinates (3-10). Actually, we can gain more information, especially in the case when $\gamma = 1$. To see this let us rewrite the Euler equation in w-coordinates by using (3-14):

$$\begin{aligned}\frac{\rho}{\alpha_1}\phi^w(w) &= \Big(v_1+w_1-\frac{1}{B}w_2\Big)\partial_1\phi^w(w)+\Big(-v_1B+A(\gamma-1)\Big(w_1-\frac{1}{B}w_2\Big)+\gamma w_2\Big)\Big(\partial_2+\frac{1}{B}\partial_1\Big)\phi^w(w)\\ &= \Big(\frac{B+A(\gamma-1)}{B}w_1+\frac{(B-A)(\gamma-1)}{B^2}w_2\Big)\partial_1\phi^w(w)\\ &\qquad\qquad +\Big(-v_1B+A(\gamma-1)w_1+\frac{B\gamma-A(\gamma-1)}{B}w_2\Big)\partial_2\phi^w(w).\end{aligned}$$

Case $\gamma = 1$. Here the Euler equation reduces to

$$\frac{\rho}{\alpha_1}\phi^w(w) = w_1\partial_1\phi^w(w) + (-v_1B + w_2)\partial_2\phi^w(w). \tag{3-18}$$

Taking the $\partial^\tau = \partial_1^{\tau_1}\partial_2^{\tau_2}$-derivative and evaluating at 0 one gets

$$\frac{\rho}{\alpha_1}\partial^\tau\phi^w(0) = \tau_1\partial^\tau\phi^w(0) - v_1B\partial_1^{\tau_1}\partial_2^{\tau_2+1}\phi^w(0) + \tau_2\partial^\tau\phi^w(0),$$

which can be rewritten as

$$\Big(\frac{\rho}{\alpha_1} - |\tau|\Big)\partial^\tau\phi^w(0) = -v_1B\partial_1^{\tau_1}\partial_2^{\tau_2+1}\phi^w(0).$$

From this and the fact from (3-17) that $\partial^\tau\phi^w(0) = 0$ for all τ satisfying $|\tau| = \tau_1 + \tau_2 \geq 2$, $0 \leq \tau_1 \leq \tilde{k} - 1$, and $0 \leq \tau_2 \leq 1$, one easily gets by induction on τ_2 that

$$\partial_1^{\tau_1}\partial_2^{\tau_2}\phi^w(0) = 0, \quad \text{when } |\tau| = \tau_1 + \tau_2 \geq 2,\ 1 \leq \tau_1 \leq \tilde{k} - 1. \tag{3-19}$$

We now prove a stronger claim, namely that

$$\begin{aligned}&\partial_1^{\tau_1}\phi^w(0, w_2) \equiv 0 \quad \text{for } 2 \leq \tau_1 \leq \tilde{k} - 1,\\ &\partial_1\phi^w(0, w_2) \equiv \partial_1\phi^w(0).\end{aligned} \tag{3-20}$$

In order to obtain this we take the $\partial_1^{\tau_1}$-derivative in (3-18) and evaluate it at $(0, w_2)$ to get

$$\Big(\frac{\rho}{\alpha_1} - \tau_1\Big)\partial_1^{\tau_1}\phi^w(0, w_2) = (-v_1B + w_2)\partial_2\partial_1^{\tau_1}\phi^w(0, w_2).$$

We note that this is a simple ordinary differential equation in w_2 of first order. It has a unique solution for $2 \leq \tau_1 \leq \tilde{k} - 1$ since $-v_1B + w_2 \neq 0$ for small w_2, and since we can take (3-19) as initial conditions. The claim for $2 \leq \tau_1 \leq \tilde{k} - 1$ follows since $\partial_1^{\tau_1}\phi^w(0, w_2) \equiv 0$ is obviously a solution. For $\tau_1 = 1$ we note that the case $\rho/\alpha_1 - \tau_1 = 0$ is trivial, and the solution is a unique constant function (necessarily equal to $\partial_1\phi^w(0)$). When $\tau_1 = 1$ and $\rho/\alpha_1 - \tau_1 \neq 0$, the differential equation evaluated at $w_2 = 0$ gives us that $\partial_1\partial_2\phi^w(0) = 0$ implies $\partial_1\phi^w(0) = 0$, which again means that $\partial_1\phi^w(0, w_2) \equiv 0$ is the unique solution of the given differential equation. We have thus proven (3-20).

Now by using Taylor approximation in w_1 for a fixed w_2, and the just-proven fact for the mapping $w_2 \mapsto \partial_1^{\tau_1}\phi^w(0, w_2)$ for $1 \le \tau_1 \le \tilde{k}-1$, we obtain that the normal form of ϕ^w (3-17) in the case $\gamma = 1$ can be rewritten as

$$\phi_v^w(w) = w_1^{\tilde{k}} r_1(w) + w_2^2 r_2(w_2),$$

where $r_1(0), r_2(0) \neq 0$. Note that now r_2 depends only on w_2. This corresponds to normal form (ii.w) when $\tilde{k}$ is finite and to normal form (i.w1) otherwise.

Case $\gamma \neq 1$. In this case we use our assumption that $A \neq 0$ in a critical way. Here it will be important to know what happens with $\partial_1^{\tau_1}\partial_2^2\phi^w(0)$ for $0 \le \tau_1 \le \tilde{k}-1$, and also how one can rewrite the normal form of the Hessian determinant $\mathcal{H}_{\phi^w}$ (and in particular its root).

Let us begin by taking the $\partial_1^{\tau_1}\partial_2$-derivative of the Euler equation in w-coordinates and evaluating it at $w = 0$. One gets

$$\begin{aligned}\frac{\rho}{\alpha_1}\partial_1^{\tau_1}\partial_2\phi^w(0) = \tau_1\frac{B + A(\gamma-1)}{B}\partial_1^{\tau_1}\partial_2\phi^w(0) + \frac{(B-A)(\gamma-1)}{B^2}\partial_1^{\tau_1+1}\phi^w(0)\\ - v_1 B\partial_1^{\tau_1}\partial_2^2\phi^w(0) + \tau_1 A(\gamma-1)\partial_1^{\tau_1-1}\partial_2^2\phi^w(0) + \frac{B\gamma - A(\gamma-1)}{B}\partial_1^{\tau_1}\partial_2\phi^w(0).\end{aligned}$$

Now recall again from (3-17) that $\partial^\tau\phi^w(0) = 0$ holds for any τ satisfying $|\tau| = \tau_1+\tau_2 \ge 2$, $0 \le \tau_1 \le \tilde{k}-1$, and $0 \le \tau_2 \le 1$. Thus, if $1 \le \tau_1 \le \tilde{k}-2$ then we get

$$v_1 B\partial_1^{\tau_1}\partial_2^2\phi^w(0) = \tau_1 A(\gamma-1)\partial_1^{\tau_1-1}\partial_2^2\phi^w(0), \tag{3-21}$$

and if $\tau_1 = \tilde{k}-1$, then

$$v_1 B\partial_1^{\tilde{k}-1}\partial_2^2\phi^w(0) = \frac{(B-A)(\gamma-1)}{B^2}\partial_1^{\tilde{k}}\phi^w(0) + (\tilde{k}-1)A(\gamma-1)\partial_1^{\tilde{k}-2}\partial_2^2\phi^w(0);$$

i.e., since $B - A = -\gamma v_2 v_1^{-1}$, we can rewrite this as

$$v_1 B\partial_1^{\tilde{k}-1}\partial_2^2\phi^w(0) + \frac{v_2\gamma(\gamma-1)}{v_1 B^2}\partial_1^{\tilde{k}}\phi^w(0) = (\tilde{k}-1)A(\gamma-1)\partial_1^{\tilde{k}-2}\partial_2^2\phi^w(0). \tag{3-22}$$

Now since $A, B, v_1 \neq 0$, and $\gamma \neq 1$, from (3-21) we may conclude by induction on τ_1 that for $0 \le \tau_1 \le \tilde{k}-2$ one has

$$\partial_1^{\tau_1}\partial_2^2\phi^w(0) \neq 0.$$

In order to unravel what is happening with $\partial_1^{\tilde{k}-1}\partial_2^2\phi^w(0)$ we need to investigate the root of $\mathcal{H}_{\phi^w}$. For this we want to solve the equation

$$x_2 - t_0 x_1^\gamma = y_2 - \binom{\gamma}{2} v_1^{-2} v_2 y_1^2 + \mathcal{O}(y_1^3) = 0,$$

in the w-coordinates, representing the homogeneity curve through v. Recall that by (3-11) we have $y_1 = w_1 - w_2/B$, $y_2 = Bw_1$, and so we want to solve

$$Bw_1 - \binom{\gamma}{2} v_1^{-2} v_2\Big(w_1 - \frac{1}{B}w_2\Big)^2 + \mathcal{O}\Big(\Big(w_1 - \frac{1}{B}w_2\Big)^3\Big) = 0$$

for the w_1-variable in terms of the w_2-variable when $|w_1|$, $|w_2|$ are small numbers. Using the above equation one gets by a simple calculation that

$$w_1 = \frac{v_2\gamma(\gamma-1)}{2v_1^2B^3}w_2^2 + \mathcal{O}(w_2^3) = w_2^2\tilde{\omega}(w_2), \tag{3-23}$$

and $\tilde{\omega} \equiv 0$ if and only if $v_2 = 0 = t_0$. Note that we have the precise value of $\tilde{\omega}(0)$. Using this we can now write down the normal form of w as

$$\begin{aligned}\phi_v^w(w) &= w_1^{\tilde{k}}r_1(w) + w_2^2r_2(w)\\ &= (w_1 - w_2^2\tilde{\omega}(w_2))^{\tilde{k}}r_1(w) + w_2^2\left(r_2(w) + \tilde{k}w_1^{\tilde{k}-1}\frac{v_2\gamma(\gamma-1)}{2v_1^2B^3}r_1(w)\right) + \mathcal{O}(w_2^4)\\ &= (w_1 - w_2^2\tilde{\omega}(w_2))^{\tilde{k}}\tilde{r}_1(w) + w_2^2\tilde{r}_2(w),\end{aligned} \tag{3-24}$$

where one can easily check by using (3-21), (3-22), (3-23), and (3-24) that $\partial_1^{\tau_1}\tilde{r}_2(0) \neq 0$ for all $0 \leq \tau_1 \leq \tilde{k}-1$, and that in fact one has the relations

$$v_1B\partial_1^{\tau_1}\tilde{r}_2(0) = \tau_1A(\gamma-1)\partial_1^{\tau_1-1}\tilde{r}_2(0)$$

for $1 \leq \tau_1 \leq \tilde{k}-1$. If $\tilde{k} = \infty$, then the above normal form in (3-24) corresponds to normal form (i.w2). Otherwise we have $3 \leq \tilde{k} < \infty$ and two subcases. Namely, if $t_0 \neq 0$ (i.e., $\tilde{\omega}(0) \neq 0$), then the above normal form corresponds to normal form (v), and if $t_0 = 0$ (and therefore $\tilde{\omega} \equiv 0$), then it corresponds to normal form (iii).

Determining $\tilde{k}$ in the special case when $\rho = \pm 1$ and $\alpha_1 = \alpha_2 \neq \rho$. According to the last line of the corresponding table for this case in Section 3D here we may assume $b_0, b_1 \neq 0$, and note that here $\gamma = 1$. We prove that the Hessian determinant of ϕ vanishes at v if and only if

$$b_2 = (1-\rho\alpha_1)\frac{b_1^2}{b_0} = \left(1 - \frac{\alpha_1}{\rho}\right)\frac{b_1^2}{b_0}. \tag{3-25}$$

In this case we furthermore have that if $\tilde{k} < \infty$ (corresponding to the case (ii.w)), then

$$\begin{aligned}b_j &= (\rho\alpha_1)^j j!\binom{\rho/\alpha_1}{j}\frac{b_1^j}{b_0^{j-1}}, \quad \text{for } j = 2, \dots, \tilde{k}-1,\\ b_{\tilde{k}} &\neq (\rho\alpha_1)^{\tilde{k}}\tilde{k}!\binom{\rho/\alpha_1}{\tilde{k}}\frac{b_1^{\tilde{k}}}{b_0^{\tilde{k}-1}},\end{aligned} \tag{3-26}$$

and if $\tilde{k} = \infty$ (corresponding to the case (i.w1)), then

$$b_j = (\rho\alpha_1)^j j!\binom{\rho/\alpha_1}{j}\frac{b_1^j}{b_0^{j-1}} \quad \text{for } j \in \{2, 3, \dots\}.$$

These formulae have already been shown for homogeneous polynomials in [Ferreyra et al. 2004, Lemma 2.2]. Therefore, we only sketch how one can prove them in our slightly more general case.

Recall from (3-9) that we have the formal series for ϕ at $y=0$:

$$\begin{aligned}\phi^y(y) &\approx (v_1+y_1)^{\rho/\alpha_1}b_0 + y_2(v_1+y_1)^{\rho/\alpha_1-1}b_1 + \frac{1}{2!}(v_1+y_1)^{\rho/\alpha_1-2}y_2^2 b_2+\cdots\\ &= \sum_{j=0}^{\infty}\frac{b_j}{j!}(v_1+y_1)^{\rho/\alpha_1-j}y_2^j.\end{aligned}$$

From this one gets

$$\partial_1^2\phi^y(0) = b_0\frac{\rho}{\alpha_1}\left(\frac{\rho}{\alpha_1}-1\right)v_1^{\rho/\alpha_1-2},\quad \partial_1\partial_2\phi^y(0)=b_1\left(\frac{\rho}{\alpha_1}-1\right)v_1^{\rho/\alpha_1-2},\quad \partial_2^2\phi^y(0)=b_2v_1^{\rho/\alpha_1-2},$$

and (3-25) follows by a direct computation (recall that $\mathcal{H}_{\phi^y}(0)=0$ if and only if $\mathcal{H}_\phi(v)=0$). More generally, we have

$$\partial^\tau\phi^y(0)=\tau_1!\binom{\rho/\alpha_1-\tau_2}{\tau_1}v_1^{\rho/\alpha_1-|\tau|}b_{\tau_2}. \tag{3-27}$$

Let us now determine the relation between y and z when the Hessian determinant vanishes. We may write

$$\begin{aligned} z_1 &= y_1, & \partial_{z_1} &= \partial_{y_1}+B\partial_{y_2},\\ z_2 &= y_2-By_1, & \partial_{z_2} &= \partial_{y_2}.\end{aligned}$$

Then by (3-25) one gets that $\partial_1^2\phi^z(0)=\partial_1\partial_2\phi^z(0)=0$ if and only if

$$B=-\frac{b_0}{b_1}\frac{\rho}{\alpha_1}.$$

From this we can determine the constant A since it is equal to t_0+B, i.e., $A=v_2/v_1-(\rho b_0)/(\alpha_1 b_1)$.

One can now directly prove (3-26) by induction in j by using (3-27), and the fact that $\partial_1^j\phi^z(0)=0$ for $2\le j<\tilde{k}$ and $\partial_1^{\tilde{k}}\phi^z(0)\neq 0$ is equivalent to

$$\begin{aligned}&\left(\partial_1-\frac{b_0}{b_1}\frac{\rho}{\alpha_1}\partial_2\right)^j\phi^y(0)=0,\quad j=2,\dots,\tilde{k}-1,\\ &\left(\partial_1-\frac{b_0}{b_1}\frac{\rho}{\alpha_1}\partial_2\right)^{\tilde{k}}\phi^y(0)\neq 0.\end{aligned}$$

We have already checked the induction base $j=2$.

3F. ***Order of vanishing of the Hessian determinant.*** In this subsection we determine the normal forms of the Hessian determinant of ϕ (or more precisely, the order of vanishing of the Hessian determinant of ϕ), as listed in Section 3A. We recall from Section 3A that if $v_1>0$, then one can write

$$\mathcal{H}_\phi(x)=(x_2-t_0x_1^\gamma)^N r_0(x),$$

where either r_0 is flat in v (which we consider as the case $N=\infty$), or $r_0(v)\neq 0$ and $0\le N<\infty$. It remains to determine N from the information provided by the normal forms of ϕ. We note that

$$N=\min\{j\ge 0:(\partial_2^j\mathcal{H})(v)\neq 0\}.$$

***Normal form* (i.y1).** First we note by the normal form tables above that this normal form appears only in cases when either $\gamma = 1$ or $t_0 = v_2 = 0$, and so we have $\omega \equiv 0$. Thus, by (3-9) the function ϕ_v^y has the formal expansion

$$\begin{aligned} \phi_v^y(y) &= \frac{1}{k!} y_2^k (y_1 + v_1)^{\rho/\alpha_1 - k\gamma} g_k(y_2(y_1+v_1)^{-1} + t_0) \\ &\approx \sum_{j=k}^{\infty} \frac{b_j}{j!} y_2^j (y_1+v_1)^{\rho/\alpha_1 - j\gamma}, \end{aligned} \tag{3-28}$$

and the Hessian determinant vanishes along $y_2 = 0$, which means we need to determine what is the least N such that $(\partial_2^N \mathcal{H}_{\phi^y})(0) \neq 0$. From the above expansion one obtains

$$\begin{aligned} &\partial_1^{\tau_1}\partial_2^{\tau_2}\phi^y(0) = 0, \quad |\tau| = \tau_1 + \tau_2 \geq 2,\ 0 \leq \tau_2 \leq k-1, \\ &\partial^\tau \phi^y(0) = \tau_1! \binom{\rho/\alpha_1 - \gamma\tau_2}{\tau_1} v_1^{\rho/\alpha_1 - \tau_1 - \gamma\tau_2} b_{\tau_2}, \quad \tau_2 \geq k. \end{aligned} \tag{3-29}$$

By applying the general Leibniz rule to the definition of the Hessian determinant we get

$$\begin{aligned} \partial_2^N \mathcal{H}_{\phi^y} &= \partial_2^N (\partial_1^2\phi^y \partial_2^2\phi^y - (\partial_1\partial_2\phi^y)^2) \\ &= \sum_{n=0}^{N} \binom{N}{n} (\partial_1^2\partial_2^n\phi^y\ \partial_2^{N+2-n}\phi^y - \partial_1\partial_2^{n+1}\phi^y\ \partial_1\partial_2^{N+1-n}\phi^y), \end{aligned} \tag{3-30}$$

and one can easily check by using (3-29) that $\partial_2^N \mathcal{H}_{\phi^y}(0) = 0$ for $N < 2k-2$. For $N = 2k-2$ we get

$$\begin{aligned} \partial_2^{2k-2}\mathcal{H}_{\phi^y}(0) &= \binom{2k-2}{k} \partial_1^2\partial_2^k\phi^y(0)\partial_2^k\phi^y(0) - \binom{2k-2}{k-1}(\partial_1\partial_2^k\phi^y)^2(0) \\ &= \left[\binom{2k-2}{k}\left(\frac{\rho}{\alpha_1} - k\gamma\right)\left(\frac{\rho}{\alpha_1} - k\gamma - 1\right) - \binom{2k-2}{k-1}\left(\frac{\rho}{\alpha_1} - k\gamma\right)^2\right] b_k^2 v_1^{2\rho/\alpha_1 - 2k\gamma - 2} \\ &= \left[\frac{k-1}{k}\left(\frac{\rho}{\alpha_1} - k\gamma - 1\right) - \left(\frac{\rho}{\alpha_1} - k\gamma\right)\right]\binom{2k-2}{k-1}\left(\frac{\rho}{\alpha_1} - k\gamma\right) b_k^2 v_1^{2\rho/\alpha_1 - 2k\gamma - 2}. \end{aligned}$$

Thus, $\partial_2^{2k-2}\mathcal{H}_{\phi^y}(0) \neq 0$ if and only if

$$\frac{\rho}{\alpha_1} \notin \{k\gamma,\ k\gamma + 1 - k\}. \tag{3-31}$$

Let us now denote by $\tilde{k}$ the smallest positive integer such that $b_{k+\tilde{k}} \neq 0$; i.e., we have

$$\begin{aligned} &b_{k+j} = 0, \quad 0 < j < \tilde{k}, \\ &b_{k+\tilde{k}} \neq 0. \end{aligned}$$

Case when $\rho/\alpha_1 = k\gamma$. By examining the term $j = k$ in (3-28) we note that in this case we additionally have

$$\begin{aligned} &\partial_2^k\phi^y(0) \neq 0 \\ &\partial_1^{\tau_1}\partial_2^k\phi^y(0) = 0, \quad \tau_1 \geq 1. \end{aligned}$$

Now by using the information in (3-29), the above additional assumption that $b_{k+j}=0$ for $0<j<\tilde{k}$, $b_{k+\tilde{k}}\neq 0$, and the Leibniz formula (3-30) a straightforward calculation yields that $\partial_2^N\mathcal{H}_{\phi^y}(0)=0$ for $N<2k+\tilde{k}-2$ and $\partial_2^{2k+\tilde{k}-2}\mathcal{H}_{\phi^y}(0)\neq 0$; i.e., we have the precise order of vanishing of the Hessian determinant.

Case when $\rho/\alpha_1=k\gamma+1-k$. Again, by a straightforward calculation using the Leibniz formula one gets that $\partial_2^N\mathcal{H}_{\phi^y}(0)=0$ for $N<2k+\tilde{k}-2$ and we have for $N=2k+\tilde{k}-2$

$$\begin{aligned}\partial_2^{2k+\tilde{k}-2}\mathcal{H}_{\phi^y}(0)&=\binom{2k+\tilde{k}-2}{k}\partial_1^2\partial_2^k\phi^y(0)\partial_2^{k+\tilde{k}}\phi^y(0)\\&\quad+\binom{2k+\tilde{k}-2}{k-2}\partial_1^2\partial_2^{k+\tilde{k}}\phi^y(0)\partial_2^k\phi^y(0)-2\binom{2k+\tilde{k}-2}{k-1}\partial_1\partial_2^k\phi^y(0)\partial_1\partial_2^{k+\tilde{k}}\phi^y(0).\end{aligned}$$

Thus

$$\begin{aligned}\left(\binom{2k+\tilde{k}-2}{k-2}\right)^{-1}\partial_2^{2k+\tilde{k}-2}\mathcal{H}_{\phi^y}(0)&=\frac{(k+\tilde{k})(k+\tilde{k}-1)}{(k-1)k}\partial_1^2\partial_2^k\phi^y(0)\partial_2^{k+\tilde{k}}\phi^y(0)\\&\quad+\partial_1^2\partial_2^{k+\tilde{k}}\phi^y(0)\partial_2^k\phi^y(0)-\frac{2(k+\tilde{k})}{k-1}\partial_1\partial_2^k\phi^y(0)\,\partial_1\partial_2^{k+\tilde{k}}\phi^y(0).\end{aligned}$$

This is equal to zero when the expression

$$\begin{aligned}&(k+\tilde{k})(k+\tilde{k}-1)\partial_1^2\partial_2^k\phi^y(0)\partial_2^{k+\tilde{k}}\phi^y(0)\\&\quad+(k-1)k\,\partial_1^2\partial_2^{k+\tilde{k}}\phi^y(0)\partial_2^k\phi^y(0)-2k(k+\tilde{k})\partial_1\partial_2^k\phi^y(0)\,\partial_1\partial_2^{k+\tilde{k}}\phi^y(0)\end{aligned}$$

equals zero. Plugging in the values of the derivatives from (3-29) one obtains that the above expression is equal to

$$(k+\tilde{k})(k+\tilde{k}-1)(1-k)(-k)+(k-1)k(1-k-\tilde{k}\gamma)(-k-\tilde{k}\gamma)-2k(k+\tilde{k})(1-k)(1-k-\tilde{k}\gamma),$$

up to the nonzero constant factor $v_1^{-2k-\tilde{k}\gamma}b_kb_{k+\tilde{k}}$. Factoring out $(1-k)(-k)$ we get

$$(k+\tilde{k})(k+\tilde{k}-1)+(k+\tilde{k}\gamma-1)(k+\tilde{k}\gamma)-2(k+\tilde{k})(k+\tilde{k}\gamma-1)$$

and this equals zero if and only if $\gamma\in\{1,(\tilde{k}+1)/\tilde{k}\}$.

The condition $\rho/\alpha_1=k\gamma+1-k$ tells us that if $\gamma=1$ then $\rho=\alpha_1=\alpha_2=1$, and from the normal form tables we see that this is precisely when the Hessian determinant vanishes to infinite order.

In the case $\gamma=(\tilde{k}+1)/\tilde{k}$ we get that $\rho=1$, $\alpha_1=\tilde{k}/(k+\tilde{k})$ and $\alpha_2=(\tilde{k}+1)/(k+\tilde{k})$. Here the order of vanishing of the Hessian determinant depends explicitly on the values b_j, and so, in contrast to the previous cases, one cannot relate in an easy way the order of vanishing of the Hessian determinant and the form of ϕ in (3-28). As we shall not need the precise order of vanishing of the Hessian determinant in this case, we do not pursue this question further.

Other normal forms. First we recall that normal form (i.y2) was dealt with in Section 3B, and there it was already determined that the Hessian vanishes of infinite order (i.e., it is flat).

In all the remaining normal forms we use either y- or w-coordinates, and so (as already noted in Section 3A) the Hessian determinant in these coordinates has the normal form

$$\mathcal{H}_{\phi^u}(u) = (u_2 - u_1^2\psi(u_1))^N r_0(u),$$

where u can represent either y- or w-coordinates, and where either N is finite and $r_0(0) \neq 0$, or the Hessian determinant is flat (in which case we consider N to be infinite). The function ψ is equal to either ω or $\tilde{\omega}$. Our goal is to determine $N = \min\{j \geq 0 : (\partial_2^j \mathcal{H}_{\phi^u})(0) \neq 0\}$.

We first note that we can rewrite all the remaining normal forms as either

$$\phi_v^u(u) = (u_2 - u_1^2\psi(u_1))^{k_0} r(u) \tag{3-32}$$

or

$$\phi_v^u(u) = u_1^2 r_1(u) + u_2^{k_0} r_2(u), \tag{3-33}$$

where $r(0), \psi(0), r_1(0), r_2(0) \neq 0$, and $k_0 \geq 2$ in the first case and $k_0 \geq 3$ in the second. In the second case $k_0 = \infty$ is allowed with an obvious interpretation. Note that the second case (3-33) includes normal forms (ii), (iii), (iv), (v), and also subcases of (i) where the w-coordinates are used. Also note that this is slightly different compared to the three forms mentioned before the detailed table of normal forms in Section 3A.

For both cases (3-32) and (3-33) one can use the Leibniz rule (3-30) and the information on the Taylor series of ϕ_v^u gained from these normal forms to obtain the order of vanishing of the Hessian determinant (in the ∂_{u_2}-direction) by a direct calculation. In the first case (3-32) one gets that the order of vanishing is $N = 2k_0 - 3$ and in the second case (3-33) one gets that $N = k_0 - 2$ (or that the Hessian determinant is flat if $k_0 = \infty$).

4. Fourier restriction when a mitigating factor is present

In this section we prove Theorem 1.1, i.e., the Fourier restriction estimate

$$\|\hat{f}\|_{L^2(\mathrm{d}\mu)} \leq C \|f\|_{L^{\mathfrak{p}_3}_{x_3}(L^{\mathfrak{p}_1}_{(x_1,x_2)})},$$

where μ is the surface measure

$$\langle \mu, f \rangle = \int_{\mathbb{R}^2 \setminus \{0\}} f(x, \phi(x))\, |\mathcal{H}_\phi(x)|^\sigma \,\mathrm{d}x$$

and the exponents are

$$\left(\frac{1}{\mathfrak{p}_1'}, \frac{1}{\mathfrak{p}_3'}\right) = \left(\frac{1}{2} - \sigma, \sigma\right).$$

The gothic letters are used in order to distinguish the endpoint exponents from the dummy ones. We assume $0 \leq \sigma < \frac{1}{2}$ when only adapted normal forms appear, and $0 \leq \sigma \leq \frac{1}{3}$ if a nonadapted normal form appears. Since the case $\sigma = 0$ follows directly by Plancherel, we may assume $\sigma > 0$.

Our assumptions in this case are that the Hessian determinant $\mathcal{H}_\phi$ does not vanish of infinite order anywhere (i.e., condition (H2) is satisfied). According to Section 2B we may restrict our attention to the localized measure

$$\langle \mu_{0,v}, f\rangle = \int_{\mathbb{R}^2\setminus\{0\}} f(x, \phi_v(x-v))\eta_v(x)|\mathcal{H}_\phi(x)|^\sigma \, dx,$$

where $v = (v_1, v_2)$ satisfies $v_1 \sim 1$, and either $v_2 = 0$ or $v_2 \sim 1$, and where η_v is a smooth nonnegative function with support in a small neighborhood of v.

After changing to y- or w-coordinates from Section 3 we get that $\mu_{0,v}$ can be rewritten as

$$\langle \nu, f\rangle = \int f(x, \phi_{\mathrm{loc}}(x))\, a(x)\, |\mathcal{H}_{\phi_{\mathrm{loc}}}(x)|^\sigma \, dx,$$

where now a is smooth, nonnegative, and supported in a small neighborhood of the origin, and where we have for ϕ_{loc} the normal form cases (i)–(vi) from Proposition 1.4. Recall that since we assume (H2), in case (i) of Proposition 1.4 the function φ vanishes identically.

The strategy will be to appropriately localize and rescale the problem, and then to use the associated "R^*R" operator. Let us begin by proving modifications of two essentially known results.

Lemma 4.1. *Let $\phi : \Omega \to \mathbb{R}$ be a smooth function on an open set $\Omega \subseteq \mathbb{R}^2$ contained in a ball of radius $\lesssim 1$, and let $\mathcal{H}_\phi = \partial_1^2\phi\partial_2^2\phi - (\partial_1\partial_2\phi)^2$ denote the Hessian determinant of ϕ. We consider the measure defined by*

$$\langle \mu, f\rangle := \int f(x_1, x_2, \phi(x))\, a(x)\, dx,$$

where $a \in C_c^\infty(\Omega)$ satisfies $\|\partial^\tau a\|_{L^\infty(\Omega)} \lesssim_\tau 1$ for all multiindices τ. If we assume that on Ω we have $|\partial_1^2\phi| \sim 1$, $|\partial^\tau\phi| \lesssim_\tau 1$ for all multiindices τ, and that $|\mathcal{H}_\phi| \sim \varepsilon$ for a bounded, strictly positive (but possibly small) constant ε, then

$$|\hat{\mu}(\xi)| \lesssim \varepsilon^{-1/2}(1+|\xi|)^{-1}.$$

The claim also holds if ϕ and a depend on ε, assuming that the implicit constants appearing in the lemma can be taken to be independent of ε.

Proof. By compactness and translating we may assume that a is supported on a small neighborhood of the origin. We also assume for simplicity that $|\partial_1\phi| \sim 1$, which can be achieved by applying a linear transformation to μ. The Fourier transform of μ is by definition

$$\hat{\mu}(\xi) = \int e^{-i\Phi(x,\xi)}\, a(x)\, dx,$$

where the phase function is of the form

$$\Phi(x,\xi) = x_1\xi_1 + x_2\xi_2 + \phi(x)\xi_3,$$

from which one easily sees that unless $|\xi_1| \sim |\xi_3| \gtrsim |\xi_2|$, we have very fast decay independent of ε. Let us define

$$s_1 = \frac{\xi_1}{\xi_3}, \quad s_2 = \frac{\xi_2}{\xi_3}, \quad \lambda = \xi_3,$$

and rewrite the phase as

$$\Phi(x,\xi) = \lambda(s_1x_1 + s_2x_2 + \phi(x)),$$

where now $|s_1| \sim 1$ and $|s_2| \lesssim 1$.

Now either the x_1-derivative of Φ has no zeros on the domain of integration (e.g., when s_1 and $\partial_1\phi(0)$ are of the same sign), in which case we get a fast decay by integrating by parts, or there is a unique zero $x_1^c = x_1^c(x_2; s_1, s_2)$ of the equation $\partial_1\Phi(x,\xi) = 0$ in x_1, depending smoothly on $(x_2; s_1, s_2)$ by the implicit function theorem, i.e., we have the relation

$$s_1 + (\partial_1\phi)(x_1^c, x_2) = 0. \tag{4-1}$$

In this case we apply the stationary phase method and get that

$$\hat{\mu}(\xi) = \lambda^{-1/2} \int e^{-i\lambda\Psi(x_2;s_1,s_2)}\, a(x_2, s_1, s_2; \lambda)\,\mathrm{d}x_2,$$

where a is a smooth function in (x_2, s_1, s_2) and a classical symbol of order 0 in λ, and where

$$\Psi(x_2; s_1, s_2) := s_1 x_1^c + s_2 x_2 + \phi(x_1^c, x_2) = \lambda^{-1}\Phi(x_1^c, x_2, \xi).$$

Taking the x_2-derivative of (4-1) we get that

$$\partial_{x_2} x_1^c(x_2; s_1, s_2) = -\frac{\partial_1\partial_2\phi(x_1^c, x_2)}{\partial_1^2\phi(x_1^c, x_2)},$$

and the x_2-derivative of the new phase is by (4-1):

$$\begin{aligned}\lambda\partial_{x_2}\Psi(x_2; s_1, s_2) &= \lambda(s_1\partial_{x_2}x_1^c + s_2 + \partial_{x_2}x_1^c\,\partial_1\phi(x_1^c, x_2) + \partial_2\phi(x_1^c, x_2)) \\ &= \lambda(s_2 + \partial_2\phi(x_1^c, x_2)).\end{aligned}$$

From this and the expression for $(x_1^c)'$ it follows that

$$\lambda\partial_{x_2}^2\Psi(x_2; s_1, s_2) = \lambda\frac{\mathcal{H}_\phi(x_1^c, x_2)}{\partial_1^2\phi(x_1^c, x_2)} \sim \lambda\varepsilon.$$

Thus, we may apply the van der Corput lemma, which then delivers the claim of the lemma. □

The following lemma for obtaining mixed-norm Fourier restriction estimates goes back essentially to [Ginibre and Velo 1992] (see also [Keel and Tao 1998]).

Lemma 4.2. *Assume that we are given a bounded open set $\Omega \subseteq \mathbb{R}^2$ and functions $\Phi \in C^\infty(\Omega; \mathbb{R}^2)$, $\phi \in C^\infty(\Omega; \mathbb{R})$, $a \in L^\infty(\Omega)$. Let us consider the measure*

$$\langle\mu, f\rangle := \int f(\Phi(x), \phi(x))\, a(x)\,\mathrm{d}x$$

and the operator $T : f \mapsto f * \hat{\mu}$. If Φ is injective and its Jacobian is of size $|J_\Phi| \sim A_1$, then the $L^1_{x_3}(\mathbb{R}; L^2_{(x_1,x_2)}(\mathbb{R}^2)) \to L^\infty_{x_3}(\mathbb{R}; L^2_{(x_1,x_2)}(\mathbb{R}^2))$ operator norm of T is bounded (up to a universal constant) by $A_1^{-1}\|a\|_{L^\infty}$. If one has furthermore the estimate

$$|\hat{\mu}(\xi)| \leq A_2(1 + |\xi_3|)^{-1},$$

then for any $\sigma \in \left[0, \frac{1}{2}\right)$ and

$$\left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1}{2} - \sigma, \sigma\right)$$

the $L^{p_3}_{x_3}(\mathbb{R}; L^{p_1}_{(x_1,x_2)}(\mathbb{R}^2)) \to L^{p_3'}_{x_3}(\mathbb{R}; L^{p_1'}_{(x_1,x_2)}(\mathbb{R}^2))$ *operator norm of* T *is bounded* (*up to a constant depending on* σ) *by* $(A_1^{-1}\|a\|_{L^\infty})^{1-2\sigma} A_2^{2\sigma}$.

Proof. Let us first introduce the operator $T_{\xi_3} g := g * \hat{\mu}(\,\cdot\,, \xi_3)$ defined for functions g on $\mathbb{R}^2$ and a fixed $\xi_3 \in \mathbb{R}$. Note that then if one writes a function f on $\mathbb{R}^3$ as $f(\xi_1, \xi_2, \xi_3) = f(\xi', \xi_3) = f_{\xi_3}(\xi')$, then

$$Tf(\xi', \xi_3) = \int (f_{\eta_3-\xi_3} * \hat{\mu}(\,\cdot\,, \eta_3))(\xi')\,\mathrm{d}\eta_3 = \int (T_{\eta_3} f_{\eta_3-\xi_3})(\xi')\,\mathrm{d}\eta_3. \tag{4-2}$$

Now note that the $L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ norm of the convolution operator T_{η_3} is bounded by the L^∞ norm of the function $(x_1, x_2) \mapsto (\mathcal{F}^{-1}_{(x_1,x_2)}\hat{\mu}(\,\cdot\,, \eta_3))(x_1, x_2)$, where for functions on $\mathbb{R}^3$ we denote by $\mathcal{F}^{-1}_{(x_1,x_2)}$ the inverse Fourier transform in the first two variables. Afterwards we can estimate the $L^1 \to L^\infty$ norm of the remaining convolution operator in η_3 by the L^∞ norm in η_3 of the kernel. Thus, for the first claim it suffices to prove that the L^∞ norm of $(\mathcal{F}^{-1}_{(x_1,x_2)}\hat{\mu})(x_1, x_2, \eta_3)$ in all three variables is bounded by $A_1^{-1}\|a\|_{L^\infty}$. In order to obtain this estimate note that $\mathcal{F}^{-1}_{(x_1,x_2)}\hat{\mu}$ is equal by Fourier inversion to the Fourier transform of μ in the third coordinate only, i.e., the distribution given by

$$\begin{aligned}
\langle \mathcal{F}_{x_3}\mu, f\rangle &= \langle \mu, \mathcal{F}_{x_3} f\rangle \\
&= \int (\mathcal{F}_{x_3} f)(\Phi(y), \phi(y))\, a(y)\,\mathrm{d}y \\
&= \iint e^{-i\eta_3\phi(y)}\, f(\Phi(y), \eta_3)\, a(y)\,\mathrm{d}\eta_3\,\mathrm{d}y \\
&= \iint e^{-i\eta_3\phi\circ\Phi^{-1}(x)}\, f(x, \eta_3)\, a\circ\Phi^{-1}(x)\, |J_\Phi(x)|^{-1}\,\mathrm{d}\eta_3\,\mathrm{d}x.
\end{aligned}$$

Thus $(\mathcal{F}_{x_3}\mu)(x_1, x_2, \eta_3)$ coincides a.e. with the function

$$(x, \eta_3) \mapsto e^{-i\eta_3\phi\circ\Phi^{-1}(x)}\, a\circ\Phi^{-1}(x)\, |J_\Phi(x)|^{-1},$$

which is now obviously bounded by $A_1^{-1}\|a\|_{L^\infty}$ up to a constant.

For the second claim note that the $L^1(\mathbb{R}^2) \to L^\infty(\mathbb{R}^2)$ norm of T_{ξ_3} is bounded by $A_2(1 + |\xi_3|)^{-1}$, and as just shown the $L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ norm is bounded up to a constant by $A_1^{-1}\|a\|_{L^\infty}$. Interpolating one gets that the $L^{p_1}(\mathbb{R}^2) \to L^{p_1'}(\mathbb{R}^2)$ norm is bounded by

$$(A_1^{-1}\|a\|_{L^\infty})^{1-2\sigma} A_2^{2\sigma} (1 + |\xi_3|)^{-2\sigma}$$

for $p_1' = \left(\frac{1}{2} - \sigma\right)$ and $\sigma \in \left[0, \frac{1}{2}\right]$. For $\sigma < \frac{1}{2}$ the claim now follows by first applying this bound to the expression (4-2) and subsequently using the (weak) Young inequality in the η_3-variable. ☐

4A. ***Normal form* (i).** In this case the local form of the phase is

$$\phi_{\mathrm{loc}}(x) = x_2^k r(x),$$

where $r(0) \neq 0$ and the Hessian determinant vanishes of order $2k + k_0 - 2$ for some $k_0 \geq 0$, i.e., it has the normal form

$$\mathcal{H}_{\phi_{\mathrm{loc}}}(x) = x_2^{2k+k_0-2} r_0(x)$$

for some smooth function r_0 satisfying $r_0(0) \neq 0$.

We begin by a dyadic decomposition $\nu = \sum_{j \gg 1} \nu_j$ in x_2 followed by scaling $x_2 \mapsto 2^{-j} x_2$. Namely, for a $j \gg 1$ we define

$$\langle \nu_j, f \rangle = \int f(x, \phi_{\mathrm{loc}}(x))\, a(x)\, \chi_1(2^j x_2)\, |\mathcal{H}_{\phi_{\mathrm{loc}}}(x)|^{\sigma}\, \mathrm{d}x,$$

where $\chi_1(x_2)$ is supported where $|x_2| \sim 1$ and is such that $\sum_{j \in \mathbb{Z}} \chi_1(x_2) = 1$. Thus, by a Littlewood–Paley argument it suffices to prove

$$\|\hat{f}\|^2_{L^2(\mathrm{d}\nu_j)} \lesssim \|f\|^2_{L^{\mathfrak{p}_3}_{x_3}(L^{\mathfrak{p}_1}_{(x_1,x_2)})}, \tag{4-3}$$

with the implicit constant independent of j. We rescale this as follows. First we note that by substituting $x_2 \mapsto 2^{-j} x_2$ we have

$$\begin{aligned}
\|\hat{f}\|^2_{L^2(\mathrm{d}\nu_j)} &= \int |\mathcal{F} f|^2(x, \phi_{\mathrm{loc}}(x))\, a(x)\, \chi_1(2^j x_2)\, |x_2|^{\sigma(2k+k_0-2)}\, |r_0(x)|^{\sigma}\, \mathrm{d}x \\
&= 2^{-j-j\sigma(2k+k_0-2)} \int |\mathcal{F} f|^2(x_1, 2^{-j} x_2, 2^{-jk} 2^{jk} \phi_{\mathrm{loc}}(x_1, 2^{-j} x_2)) \\
&\qquad\qquad \times a(x_1, 2^{-j} x_2)\, \chi_1(x_2)\, |x_2|^{\sigma(2k+k_0-2)}\, |r_0(x_1, 2^{-j} x_2)|^{\sigma}\, \mathrm{d}x \\
&= 2^{-j-j\sigma(2k+k_0-2)} \int |\mathcal{F} f|^2(x_1, 2^{-j} x_2, 2^{-jk} \tilde{\phi}(x, 2^{-j})))\, a(x, 2^{-j})\, \mathrm{d}x.
\end{aligned}$$

The last expression can be rewritten as

$$2^{-j-j\sigma(2k+k_0-2)} \langle \tilde{\nu}_j, |\operatorname{Dil}_{(1,2^j,2^{jk})}(\mathcal{F} f)|^2 \rangle,$$

where

$$\langle \tilde{\nu}_j, f \rangle = \int f(x, \tilde{\phi}(x, 2^{-j}))\, a(x, 2^{-j})\, \mathrm{d}x.$$

The amplitude $a(x, 2^{-j})$ is now supported so that $|x_1| \ll 1$ and $|x_2| \sim 1$, and it is C^∞ having derivatives uniformly bounded. The phase is

$$\tilde{\phi}(x, 2^{-j}) = 2^{jk} \phi_{\mathrm{loc}}(x_1, 2^{-j} x_2) = x_2^k r(x_1, 2^{-j} x_2).$$

Now the inequality (4-3) can be rewritten as

$$\langle \tilde{\nu}_j, |\operatorname{Dil}_{(1,2^j,2^{jk})}(\mathcal{F} f)|^2 \rangle \lesssim 2^{j+j\sigma(2k+k_0-2)} \|f\|^2_{L^{\mathfrak{p}_3}_{x_3}(L^{\mathfrak{p}_1}_{(x_1,x_2)})}.$$

Interchanging the dilation and the Fourier transform we get

$$2^{2j+2jk} \langle \tilde{\nu}_j, |\mathcal{F}(\operatorname{Dil}_{(1,2^{-j},2^{-jk})} f)|^2 \rangle \lesssim 2^{j+j\sigma(2k+k_0-2)} \|f\|^2_{L^{\mathfrak{p}_3}_{x_3}(L^{\mathfrak{p}_1}_{(x_1,x_2)})},$$

and this is equivalent to

$$\begin{aligned}
2^{2j+2jk} \langle \tilde{\nu}_j, |\hat{f})|^2 \rangle &\lesssim 2^{j+j\sigma(2k+k_0-2)} \|\operatorname{Dil}_{(1,2^j,2^{jk})} f\|^2_{L^{\mathfrak{p}_3}_{x_3}(L^{\mathfrak{p}_1}_{(x_1,x_2)})} \\
&= 2^{j+j\sigma(2k+k_0-2)} 2^{2j/\mathfrak{p}_1+2jk/\mathfrak{p}_3} \|f\|^2_{L^{\mathfrak{p}_3}_{x_3}(L^{\mathfrak{p}_1}_{(x_1,x_2)})}.
\end{aligned}$$

Plugging in the values of $\mathfrak{p}_1$ and $\mathfrak{p}_3$ we finally obtain

$$\|\hat{f}\|^2_{L^2(\mathrm{d}\tilde{\nu}_j)} \lesssim 2^{\sigma j k_0} \|f\|^2_{L^{\mathfrak{p}_3}_{x_3}(L^{\mathfrak{p}_1}_{(x_1,x_2)})}; \tag{4-4}$$

i.e., this is the rescaled form of the (4-3) inequality.

Now note that from the expression for $\tilde{\phi}(x, 2^{-j})$ we have $|\partial_2\tilde{\phi}| \sim 1 \sim |\partial_2^2\tilde{\phi}|$ and one easily gets by using the definition of the Hessian determinant that

$$\begin{aligned}\mathcal{H}_{\tilde{\phi}}(x, 2^{-j}) &= 2^{j(2k-2)}\mathcal{H}_{\phi_{\mathrm{loc}}}(x_1, 2^{-j}x_2)\\ &= 2^{-jk_0} x_2^{2k+k_0-2} r_0(x_1, 2^{-j}x_2).\end{aligned}$$

Thus $|\mathcal{H}_{\tilde{\phi}}(x, 2^{-j})| \sim 2^{-jk_0}$, from which the estimate (4-4) follows by an application of Lemma 4.1 and subsequently Lemma 4.2.

4B. ***Preliminary rescaling for cases* (ii)–(vi).** In normal form cases (ii)–(vi) the principal face of $\mathcal{N}(\phi_{\mathrm{loc}})$ is compact and so we use the scaling associated to it:

$$\delta_t^\kappa(x) = (t^{\kappa_1}x_1, t^{\kappa_2}x_2),$$

where in cases (ii)-(v) we have

$$\kappa = \left(\frac{1}{2}, \frac{1}{k}\right)$$

and in case (vi) we have

$$\kappa = \left(\frac{1}{2k}, \frac{1}{k}\right).$$

In particular, for $j \gg 1$ we define

$$\langle \nu_j, f\rangle = \int f(x, \phi_{\mathrm{loc}}(x))\, a(x)\, \eta(\delta^\kappa_{2^j}x)\, |\mathcal{H}_{\phi_{\mathrm{loc}}}(x)|^\sigma \,\mathrm{d}x,$$

where η is supported on an annulus and is such that $\sum_{j\in\mathbb{Z}} \eta(\delta^\kappa_{2^j}x) = 1$. By using Littlewood–Paley theory we get that it is sufficient to prove

$$\|\hat{f}\|^2_{L^2(\mathrm{d}\nu_j)} \lesssim \|f\|^2_{L^{\mathfrak{p}_3}_{x_3}(L^{\mathfrak{p}_1}_{(x_1,x_2)})}.$$

Rescaling similarly as in the case of normal form (i), the above estimate is equivalent to

$$\|\hat{f}\|^2_{L^2(\mathrm{d}\tilde{\nu}_j)} \lesssim \|f\|^2_{L^{\mathfrak{p}_3}_{x_3}(L^{\mathfrak{p}_1}_{(x_1,x_2)})}, \tag{4-5}$$

where

$$\langle \tilde{\nu}_j, f\rangle = \int f(x, \tilde{\phi}(x,\delta))\, |\mathcal{H}_{\tilde{\phi}}(x,\delta)|^\sigma \, a(x,\delta)\,\mathrm{d}x. \tag{4-6}$$

Here the amplitude $a(x, \delta)$ is supported on a fixed annulus around the origin,

$$\delta = (\delta_0, \delta_1, \delta_2) := (2^{-j(k-1)/k}, 2^{-j/2}, 2^{-j/k}) \tag{4-7}$$

in cases (ii)–(v), and

$$\delta = (\delta_1, \delta_2) := (2^{-j/(2k)}, 2^{-j/k})$$

in case (vi). The phase which one obtains in (4-6) is

$$\tilde{\phi}(x,\delta) := 2^j \phi_{\mathrm{loc}}(\delta_1 x_1, \delta_2 x_2).$$

The quantity δ_0 will be appear only later when we use the explicit normal forms. From the above phase form it follows that

$$\mathcal{H}_{\tilde{\phi}}(x,\delta) = 2^{j(k-2)/k} \mathcal{H}_{\phi_{\mathrm{loc}}}(\delta_1 x_1, \delta_2 x_2)$$

in cases (ii)–(v), and

$$\mathcal{H}_{\tilde{\phi}}(x,\delta) = 2^{j(2k-3)/k} \mathcal{H}_{\phi_{\mathrm{loc}}}(\delta_1 x_1, \delta_2 x_2)$$

in case (vi).

4C. *Normal forms* (ii) *and* (iii). Using the normal forms for ϕ_{loc} one gets in these cases

$$\begin{aligned}\tilde{\phi}(x,\delta) &= x_1^2 r_1(\delta_1 x_1, \delta_2 x_2) + x_2^k r_2(\delta_1 x_1, \delta_2 x_2),\\ \mathcal{H}_{\tilde{\phi}}(x,\delta) &= x_2^{k-2} r_0(\delta_1 x_1, \delta_2 x_2),\end{aligned}$$

where $r_0(0), r_1(0), r_2(0) \neq 0$, and $k \geq 3$. Hence, for the part where $|x_2| \gtrsim 1$ in (4-6) the Hessian is nondegenerate, and so we may localize to $|x_1| \sim 1$ and $|x_2| \ll 1$, and subsequently perform a dyadic decomposition in the x_2-coordinate; i.e., we define

$$\begin{aligned}\langle \nu_l, f\rangle &:= \int f(x, \tilde{\phi}(x,\delta))\, |x_2|^{\sigma(k-2)}\, \chi_1(2^l x_2)\, a(x,\delta)\, \mathrm{d}x\\ &= 2^{-l-l\sigma(k-2)} \int f(x_1, 2^{-l}x_2, \tilde{\phi}(x_1, 2^{-l}x_2, \delta))\, a(x,\delta,2^{-l})\, \mathrm{d}x,\end{aligned}$$

where now the amplitude is supported in a domain where $|x_1| \sim 1 \sim |x_2|$ and has uniformly bounded C^N norm for any N. Applying the Littlewood–Paley theorem again and rescaling, it is sufficient for us to prove

$$\|\hat{f}\|^2_{L^2(\mathrm{d}\tilde{\nu}_{j,l})} \lesssim 2^{kl\sigma} \|f\|^2_{L^{\mathfrak{p}_3}_{x_3}(L^{\mathfrak{p}_1}_{(x_1,x_2)})}, \tag{4-8}$$

where the rescaled measure is

$$\langle \tilde{\nu}_{j,l}, f\rangle = \int f(x, \tilde{\phi}(x_1, 2^{-l}x_2, \delta))\, a(x,\delta,2^{-l})\, \mathrm{d}x.$$

The phase has now the form

$$x_1^2 r_1(\delta_1 x_1, \delta_2 x_2) + 2^{-kl} x_2^k r_2(\delta_1 x_1, 2^{-l}\delta_2 x_2) \tag{4-9}$$

on the domain $|x_1| \sim 1$ and $|x_2| \sim 1$, and its Hessian determinant is of size 2^{-kl}. By Lemma 4.1 we have

$$|\hat{\tilde{\nu}}_{j,l}(\xi)| \lesssim 2^{kl/2}(1+|\xi|)^{-1}.$$

And so the estimate (4-8) follows by Lemma 4.2.

4D. ***Normal form* (iv).** In this case we get

$$\tilde{\phi}(x,\delta) = x_1^2 q(\delta_1 x_1) + (x_2 - \delta_0 x_1^2 \psi(\delta_1 x_1))^k r(\delta_1 x_1, \delta_2 x_2),$$
$$\mathcal{H}_{\tilde{\phi}}(x,\delta) = (x_2 - \delta_0 x_1^2 \psi(\delta_1 x_1))^{k-2} r_0(\delta_1 x_1, \delta_2 x_2),$$

where $q(0), r(0), r_0(0), \psi(0) \neq 0$, and $k \geq 3$. Therefore again, if $|x_2| \gtrsim 1$ the Hessian is nondegenerate and therefore we may concentrate on $|x_1| \sim 1$ and $|x_2| \ll 1$ in (4-6). We perform a dyadic decomposition, though this time depending on how close we are to the root of the Hessian determinant, i.e., we define

$$\langle \nu_l, f \rangle := \int f(x, \tilde{\phi}(x,\delta))\, |x_2 - \delta_0 x_1^2 \psi(\delta_1 x_1)|^{\sigma(k-2)}\, \chi_1(2^l (x_2 - \delta_0 x_1^2 \psi(\delta_1 x_1)))\, a(x,\delta)\, \mathrm{d}x.$$

Next, after changing coordinates from x_2 to $x_2 + \delta_0 x_1^2 \psi(\delta_1 x_1)$ we may write

$$\langle \nu_l, f \rangle = \int f(x_1, x_2 + \delta_0 x_1^2 \psi(\delta_1 x_1), \phi_1(x,\delta))\, |x_2|^{\sigma(k-2)}\, \chi_1(2^l x_2)\, a_1(x,\delta)\, \mathrm{d}x, \tag{4-10}$$

where

$$\begin{aligned} \phi_1(x,\delta) &= x_1^2 q(\delta_1 x_1) + x_2^k r(\delta_1 x_1, \delta_2 x_2 + \delta_0 \delta_2 x_1^2 \psi(\delta_1 x_1, \delta_2 x_2)) \\ &= x_1^2 q(\delta_1 x_1) + x_2^k r(\delta_1 x_1, \delta_2 x_2 + (\delta_1 x_1)^2 \psi(\delta_1 x_1, \delta_2 x_2)) \\ &= x_1^2 q(\delta_1 x_1) + x_2^k \tilde{r}(\delta_1 x_1, \delta_2 x_2). \end{aligned}$$

The function $\tilde{r}$ is a smooth and nonzero at the origin. Finally, we rescale in x_2 as $x_2 \mapsto 2^{-l} x_2$ and may write

$$\begin{aligned} &\langle \nu_l, f \rangle \\ &= 2^{-l-l\sigma(k-2)} \int f(x_1, 2^{-l} x_2 + \delta_0 x_1^2 \psi(\delta_1 x_1), \phi_{j,l}(x,\delta,2^{-l}))\chi_1(x_1)\, \chi_1(x_2)\, a(x,\delta,2^{-l})\, \mathrm{d}x, \end{aligned} \tag{4-11}$$

where the amplitude is a smooth function and the phase is

$$\phi_{j,l}(x,\delta) = x_1^2 q(\delta_1 x_1) + 2^{-kl} x_2^k \tilde{r}(\delta_1 x_1, 2^{-l} \delta_2 x_2).$$

In order to obtain the estimate (4-5) we shall need essentially a variant of Lemma 4.2. Namely, we shall consider the analytic family of operators T_ζ defined by convolution against the Fourier transform of the measure

$$\mu_\zeta := \sum_{2^l \gg 1} 2^{l\sigma(k-2)} 2^{-l\zeta(k-2)} \nu_l, \tag{4-12}$$

where ζ has real part between 0 and $\frac{1}{2}$, and in particular, for a fixed $\xi_3 \in \mathbb{R}^3$, we shall consider the operator $T_\zeta^{\xi_3} : f \mapsto f * \hat{\mu}_\zeta(\cdot, \xi_3)$. Note that we are interested in μ_σ since this is precisely the sum of measures ν_l.

When the real part of ζ is 0 (i.e., $\zeta = it,\ t \in \mathbb{R}$) one considers the $L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ estimate for which we use (4-10). In (4-10) we see that the amplitude is of size $2^{-l\sigma(k-2)}$, which is precisely what we need in (4-12). Since the supports are disjoint when varying l, we get by an argument similar to that in Lemma 4.2 that the operator $L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ norm of $T_{it}^{\xi_3}$ is $\lesssim 1$ (uniform in ξ_3 and t).

When the real part of ζ is $\frac{1}{2}$ we need to prove

$$|\hat{\mu}_{1/2+it}(\xi)| \lesssim (1 + |\xi_3|)^{-1} \tag{4-13}$$

with implicit constant independent of t and ξ_3, since this would give us that the operator norm of $T^{\xi_3}_{1/2+it}$ for mapping $L^1(\mathbb{R}^2) \to L^\infty(\mathbb{R}^2)$ is bounded by $(1+|\xi_3|)^{-1}$.

Thus, under the assumption that we have the estimate (4-13) we may apply complex interpolation for each fixed ξ_3 to the analytic family of operators $T^{\xi_3}_\zeta$ and obtain that the operator norm of $T^{\xi_3}_\sigma$ between spaces $L^{\mathfrak{p}_1}(\mathbb{R}^2) \to L^{\mathfrak{p}_1'}(\mathbb{R}^2)$ is $\lesssim (1+|\xi_3|)^{-2\sigma}$, and so in the same way as in the proof of Lemma 4.2 the (weak) Young inequality in the x_3-direction implies (4-5).

In proving (4-13) it suffices to show that

$$\sum_{2^l \gg 1} 2^{-l(1/2-\sigma)(k-2)} |\hat{\nu}_l(\xi)| \lesssim (1+|\xi_3|)^{-1}$$

for all $\xi \in \mathbb{R}^3$. By (4-11) the Fourier transform of a summand is

$$2^{-l(1/2-\sigma)(k-2)} \hat{\nu}_l(\xi) = 2^{-kl/2} \int e^{-i\Phi(x,\xi,\delta,2^{-l})}\, \chi_1(x_1)\, \chi_1(x_2)\, a(x,\delta,2^{-l})\, \mathrm{d}x,$$

where the phase function is

$$\Phi(x,\xi,\delta,2^{-l}) := \xi_1 x_1 + \xi_2 \delta_0 x_1^2 \psi(\delta_1 x_1) + \xi_3 x_1^2 q(\delta_1 x_1) + 2^{-l}\xi_2 x_2 + 2^{-kl}\xi_3 x_2^k \tilde{r}(\delta_1 x_1, 2^{-l}\delta_2 x_2).$$

We see that when either $|\xi_1| \gg \max\{|\xi_2|,|\xi_3|\}$ or $|\xi_3| \gg \max\{|\xi_1|,|\xi_2|\}$ we can use integration by parts in the x_1-variable and get a very fast decay. This is also the case when $|\xi_1| \sim |\xi_2|$ are much greater than $|\xi_3|$, or when $|\xi_2| \sim |\xi_3|$ are much greater than $|\xi_1|$. If we have $|\xi_2| \gtrsim |\xi_3|$, then we may use integration by parts in x_2 and get

$$|2^{-l(1/2-\sigma)(k-2)} \hat{\nu}_l(\xi)| \lesssim 2^{-kl/2}(1+2^{-l}|\xi_2|)^{-1} \lesssim 2^{-kl/2}(1+2^{-l}|\xi_3|)^{-1},$$

from which (4-13) follows since $k \geq 3$. We are thus left with the case when $|\xi_1| \sim |\xi_3| \gg |\xi_2|$.

Case 1: $2^{-kl}|\xi_3| \lesssim 1$. Here we use the van der Corput lemma in x_1 only and get

$$|2^{-l(1/2-\sigma)(k-2)} \hat{\nu}_l(\xi)| \lesssim 2^{-kl/2}|\xi_3|^{-1/2}.$$

Summation in l then gives precisely (4-13).

Case 2: $2^{-l}|\xi_2| \nsim 2^{-kl}|\xi_3|$ and $2^{-kl}|\xi_3| \gg 1$. We may use in this case integration by parts in x_2 and then the van der Corput lemma in x_1 and get

$$|2^{-l(1/2-\sigma)(k-2)} \hat{\nu}_l(\xi)| \lesssim 2^{-kl/2}|\xi_3|^{-1/2}\,(2^{-kl}|\xi_3|)^{-1} \lesssim 2^{kl/2}|\xi_3|^{-3/2}.$$

We may now sum in l.

Case 3: $2^{-l}|\xi_2| \sim 2^{-kl}|\xi_3| \gg 1$. Here we have by iterative stationary phase (first in x_2 and then in x_1) that

$$|2^{-l(1/2-\sigma)(k-2)} \hat{\nu}_l(\xi)| \lesssim 2^{-kl/2}|\xi_3|^{-1/2}\,(2^{-kl}|\xi_3|)^{-1/2} = |\xi_3|^{-1}.$$

Here we note that $2^{l(k-1)} \sim |\xi_3||\xi_2|^{-1}$, and so we sum only over finitely many (i.e., $\mathcal{O}(1)$) l for each fixed ξ. Thus, here we also have the estimate (4-13).

4E. ***Normal form*** **(v).** Recall that here

$$\phi_{\mathrm{loc}}(x) = x_1^2 r_1(x) + (x_2 - x_1^2\psi(x_1))^k r_2(x),$$
$$\mathcal{H}_{\phi_{\mathrm{loc}}}(x) = (x_2 - x_1^2\psi(x_1))^{k-2} r_0(x),$$

where we know that $k \geq 3$, $r_0(0), r_1(0), r_2(0), \psi(0) \neq 0$. Furthermore, recall that this corresponded to the w-coordinates when deriving the normal forms, and we have shown that we additionally have in this case

$$\partial_2^{\tau_2} r_1(0) \neq 0 \quad \text{for all } \tau_2 \in \{0, 1, \dots, k-1\}.$$

In fact, one has the relationship

$$c\tau_2 \partial_2^{\tau_2 - 1} r_1(0) = \partial_2^{\tau_2} r_1(0) \quad \text{for all } \tau_2 \in \{1, \dots, k-1\},$$

where c is some fixed nonzero constant (see Section 3E). This implies for example the relation

$$r_1(0)\, \partial_2^2 r_1(0) - 2(\partial_2 r_1)^2(0) = 0. \tag{4-14}$$

From the above normal form we have

$$\tilde{\phi}(x,\delta) = x_1^2 r_1(\delta_1 x_1, \delta_2 x_2) + (x_2 - \delta_0 x_1^2 \psi(\delta_1 x_1))^k r_2(\delta_1 x_1, \delta_2 x_2),$$
$$\mathcal{H}_{\tilde{\phi}}(x,\delta) = (x_2 - \delta_0 x_1^2 \psi(\delta_1 x_1))^{k-2} r_0(\delta_1 x_1, \delta_2 x_2).$$

We may as usual localize to $|x_1| \sim 1$ and $|x_2| \ll 1$. We shall abuse the notation a bit and denote this localized measure again by $\tilde{\nu}_j$. After changing coordinates from x_2 to $x_2 + \delta_0 x_1^2 \psi(\delta_1 x_1)$ we may write

$$\langle \tilde{\nu}_j, f \rangle = \int f(x_1, x_2 + \delta_0 x_1^2 \psi(\delta_1 x_1), \phi_1(x,\delta))\, |x_2|^{\sigma(k-2)}\, a_1(x,\delta)\, \chi_1(x_1)\, \chi_0(x_2)\, \mathrm{d}x,$$

with the phase being

$$\phi_1(x,\delta) = x_1^2 \tilde{r}_1(\delta_1 x_1, \delta_2 x_2) + x_2^k \tilde{r}_2(\delta_1 x_1, \delta_2 x_2),$$

where $\tilde{r}_1, \tilde{r}_2$ are smooth functions, nonzero at the origin, and satisfy the same properties and relations as r_1 and r_2 mentioned at the beginning of this subsection. As in the case (iv), we also decompose the measure $\tilde{\nu}_j$ as $\tilde{\nu}_j = \sum_l \nu_l$, where

$$\langle \nu_l, f \rangle = \int f(x_1, x_2 + \delta_0 x_1^2 \psi(\delta_1 x_1), \phi_1(x,\delta))\, |x_2|^{\sigma(k-2)}\, a_1(x,\delta)\, \chi_1(x_1)\, \chi_1(2^l x_2)\, \mathrm{d}x.$$

Next, we shall be interested in the rescaled phase

$$\phi_l(x,\delta,2^{-l}) = \phi_1(x_1, 2^{-l}x_2, \delta) = \tilde{\phi}(x_1, 2^{-l}x_2 + \delta_0 x_1^2 \psi(\delta_1 x_1), \delta).$$

Now we need a relation between the Hessian determinant of ϕ_l and the Hessian determinant of $\tilde{\phi}$. For this let us define for simplicity

$$\varphi(x_1, \delta_1) := \delta_1^2 x_1^2 \psi(\delta_1 x_1).$$

The reason why we have not included the factor δ_2^{-1} will be clear later (recall from (4-7) that $\delta_0 = \delta_1^2\delta_2^{-1}$). A direct calculation shows then

$$\mathcal{H}_{\phi_l} = 2^{-2l}\mathcal{H}_{\tilde{\phi}} + \delta_2^{-1}2^l\partial_1^2\varphi\partial_2\phi_l\partial_2^2\phi_l, \tag{4-15}$$

and due to our localization we have $|\mathcal{H}_{\tilde{\phi}}| \sim 2^{-l(k-2)}$.

We use the same complex interpolation idea as in (iv) according to which it suffices to prove

$$\sum_{2^l \gg 1} 2^{-l(1/2-\sigma)(k-2)}|\hat{\nu}_l(\xi)| \lesssim (1+|\xi_3|)^{-1},$$

where after rescaling $x_2 \mapsto 2^{-l}x_2$ we have

$$2^{-l(1/2-\sigma)(k-2)}\hat{\nu}_l(\xi) = 2^{-kl/2}\int e^{i\Phi_0(x,\xi,\delta,2^{-l})}a(x,\delta,2^{-l})\,dx,$$

where the phase function for the Fourier transform of ν_l is

$$\begin{aligned}\Phi_0&(x,\xi,\delta,2^{-l})\\ &:= \xi_1x_1 + \xi_2\delta_0x_1^2\psi(\delta_1x_1) + \xi_3x_1^2\tilde{r}_1(\delta_1x_1, 2^{-l}\delta_2x_2) + \xi_2 2^{-l}x_2 + \xi_3 2^{-kl}x_2^k\tilde{r}_2(\delta_1x_1, 2^{-l}\delta_2x_2)\\ &= \xi_1x_1 + \xi_2\delta_2^{-1}\varphi(x_1,\delta_1) + \xi_2 2^{-l}x_2 + \xi_3\phi_l(x,\delta,2^{-l}).\end{aligned}$$

The amplitude localizes the integration to $|x_1| \sim 1 \sim |x_2|$.

Using the same argumentation as in the case (iv) we can reduce ourselves to the case when $|\xi_1| \sim |\xi_3|$, $|\xi_2| \ll |\xi_3|$, and $|\xi_3|2^{-kl} \gg 1$ are satisfied.

Now let us make some further reductions using the fact that $\partial_2\tilde{r}_1(0), \partial_2^2\tilde{r}_1(0) \neq 0$. The x_2-derivative of the phase Φ_0 contains three terms of respective sizes $\sim |2^{-l}\delta_2\xi_3|$, $\sim |2^{-l}\xi_2|$, and $\sim |2^{-kl}\xi_3|$. If we may integrate by parts in x_2 (i.e., if one of the above terms is much larger than the other two), we can get an admissible estimate and sum in l. If $|2^{-kl}\xi_3|$ is comparable to the larger of the other two terms, then one easily sees that the second derivative in x_2 is necessarily of size $|2^{-kl}\xi_3|$, and so in this case we get by iterative stationary phase the estimate

$$2^{-l(1/2-\sigma)(k-2)}|\hat{\nu}_l(\xi)| \lesssim (1+|\xi_3|)^{-1}.$$

Note that we do not need to sum in l since there are only finitely many l satisfying one of the relations $|2^{-kl}\xi_3| \sim |2^{-l}\delta_2\xi_3|$ or $|2^{-kl}\xi_3| \sim |2^{-l}\xi_2|$.

We are thus now reduced to the case when

$$|2^{-l}\xi_2| \sim |2^{-l}\delta_2\xi_3| \gg |2^{-kl}\xi_3|, \quad |\xi_1| \sim |\xi_3| \quad \text{and} \quad |\xi_3|2^{-kl} \gg 1.$$

At this point we introduce some further notation,

$$\lambda := \xi_3, \quad s_1 := \frac{\xi_1}{\xi_3}, \quad s_2 := \frac{\xi_2}{\delta_2\xi_3}, \quad \varepsilon := 2^{-l}\delta_2,$$

and so we have $|s_1| \sim 1 \sim |s_2|$, $\lambda 2^{-kl} \gg 1$, and $\varepsilon \gg 2^{-kl}$. The phase Φ_0 can now be rewritten as $\lambda\Phi$, where Φ is

$$\Phi(x, s_1, s_2, \delta_1, \varepsilon, 2^{-kl}) = s_1 x_1 + s_2 \delta_1^2 x_1^2 \psi(\delta_1 x_1) + s_2 \varepsilon x_2 + \phi_l(x, \delta, 2^{-l}),$$

since we note from the form of ϕ_l that ϕ_l can also be taken to depend on $(x_1, x_2, \delta_1, \varepsilon, 2^{-kl})$.

Let us now apply the stationary phase method in x_1. We may rewrite the phase as

$$\Phi(x, s_1, s_2, \delta_1, \varepsilon, 2^{-kl}) = s_1 x_1 + s_2 \varphi + s_2 \varepsilon x_2 + \phi_l,$$

where we recall that $\varphi(x_1, \delta) = \delta_1^2 x_1^2 \psi(\delta_1 x_1)$. We may assume that there is a stationary point for the x_1-derivative since $|\partial_1^2 \phi_l| \sim 1$ and $|s_1| \sim 1$, and as otherwise we may use integration by parts.

We denote by $x_1^c = x_1^c(x_2, s_1, s_2, \delta_1, \varepsilon, 2^{-kl})$ the function such that

$$(\partial_1 \Phi)(x_1^c, x_2, s_1, s_2, \delta_1, \varepsilon, 2^{-kl}) = s_1 + s_2 \partial_1 \varphi + \partial_1 \phi_l = 0. \tag{4-16}$$

Taking the x_2-derivative we get

$$s_2 (x_1^c)' \partial_1^2 \varphi + (x_1^c)' \partial_1^2 \phi_l + \partial_1 \partial_2 \phi_l = 0. \tag{4-17}$$

After applying the stationary phase method in x_1 we gain a decay factor of $\lambda^{-1/2}$; i.e., we have

$$2^{-l(1/2-\sigma)(k-2)} \hat{\nu}_l(\xi) = \lambda^{-1/2} 2^{-kl/2} \int e^{-i\lambda \tilde{\Phi}(x_2, s_1, s_2, \delta_1, \varepsilon, 2^{-kl})} a(x_2, s_1, s_2, \delta, 2^{-l}; \lambda) \, \mathrm{d}x_2,$$

where the new phase is

$$\tilde{\Phi}(x_2, s_1, s_2, \delta_1, \varepsilon, 2^{-kl}) = s_1 x_1^c + s_2 \varphi(x_1^c, \delta_1) + s_2 \varepsilon x_2 + \phi_l(x_1^c, x_2, \delta, 2^{-l}),$$

and the amplitude a is a classical symbol in λ of order 0.

Taking the x_2-derivative of the expression for the new phase $\tilde{\Phi}$ and using (4-16) we get

$$\tilde{\Phi}' = s_2 \varepsilon + \partial_2 \phi_l. \tag{4-18}$$

Therefore, the second derivative of the new phase is

$$\tilde{\Phi}'' = (\partial_2 \phi_l)' = \partial_2^2 \phi_l + (x_1^c)' \partial_1 \partial_2 \phi_l. \tag{4-19}$$

Now using in order (4-17), the definition of $\mathcal{H}_{\phi_l}$ (4-15), (4-18), and (4-19), we obtain

$$\begin{aligned}
(\partial_1^2 \phi_l) \tilde{\Phi}'' &= \partial_1^2 \phi_l \, \partial_2^2 \phi_l + \partial_1 \partial_2 \phi_l (-\partial_1 \partial_2 \phi_l - s_2 (x_1^c)' \partial_1^2 \varphi) \\
&= \mathcal{H}_{\phi_l} - s_2 (x_1^c)' \partial_1^2 \varphi \, \partial_1 \partial_2 \phi_l \\
&= 2^{-2l} \mathcal{H}_{\tilde{\phi}} + \delta_2^{-1} 2^l \partial_1^2 \varphi \partial_2 \phi_l \partial_2^2 \phi_l - s_2 (x_1^c)' \partial_1^2 \varphi \, \partial_1 \partial_2 \phi_l \\
&= 2^{-2l} \mathcal{H}_{\tilde{\phi}} + \varepsilon^{-1} \partial_1^2 \varphi \, \partial_2^2 \phi_l (\tilde{\Phi}' - \varepsilon s_2) - s_2 (x_1^c)' \partial_1^2 \varphi \, \partial_1 \partial_2 \phi_l \\
&= 2^{-2l} \mathcal{H}_{\tilde{\phi}} - s_2 \partial_1^2 \varphi \, \partial_2^2 \phi_l - s_2 (x_1^c)' \partial_1^2 \varphi \, \partial_1 \partial_2 \phi_l + \varepsilon^{-1} \partial_1^2 \varphi \, \partial_2^2 \phi_l \, \tilde{\Phi}' \\
&= 2^{-2l} \mathcal{H}_{\tilde{\phi}} - s_2 \partial_1^2 \varphi \, \tilde{\Phi}'' + \varepsilon^{-1} \partial_1^2 \varphi \, \partial_2^2 \phi_l \, \tilde{\Phi}'.
\end{aligned}$$

Thus, we get

$$(s_2\partial_1^2\varphi+\partial_1^2\phi_l)\widetilde{\Phi}''=2^{-2l}\mathcal{H}_{\tilde{\phi}}+\varepsilon^{-1}\partial_1^2\varphi\,\partial_2^2\phi_l\,\widetilde{\Phi}'. \tag{4-20}$$

Note that we have $|\varepsilon^{-1}\partial_1^2\varphi\,\partial_2^2\phi_l|\ll\delta_1^2\ll 1$ and $|s_2\partial_1^2\varphi+\partial_1^2\phi_l|\sim 1$, and recall that $|2^{-2l}\mathcal{H}_{\tilde{\phi}}|\sim 2^{-kl}$. We claim that either $|\widetilde{\Phi}'|\lesssim 2^{-kl}$ on the whole domain of integration (i.e., for $|x_2|\sim 1$), or that $|\widetilde{\Phi}'|\gtrsim 2^{-kl}$ on the whole domain of integration. This can be shown by using the formula for the solution of a linear first-order ODE (considering $\widetilde{\Phi}'$ as the unknown), or by arguing by contradiction.

Let us argue by contradiction in the following way. Let us assume that there exists a point $|x_2^0|\sim 1$ such that $|\widetilde{\Phi}'(x_2^0)|\le 2^{-kl}$. Furthermore, let us assume that there exists a point $|x_2^1|\sim 1$ where one has $|\widetilde{\Phi}'|=C_1 2^{-kl}$ for some sufficiently large C_1, and let us assume that x_2^1 is the closest point to x_2^0 satisfying this condition in the sense that $|\widetilde{\Phi}'|<C_1 2^{-kl}$ between x_2^0 and x_2^1. Then the mean value theorem implies that there is a point between x_2^0 and x_2^1 where we have $|\widetilde{\Phi}''|\ge C_2 2^{-kl}$, where C_2 can be taken to tend to ∞ as C_1 tends to ∞. On the other hand, (4-20) implies that on the interval between x_2^0 and x_2^1 we have $|\widetilde{\Phi}''|\le C_3 2^{-kl}$, where we can take C_3 to be a fixed constant if δ_1 is taken to be sufficiently small when C_1 and C_2 are large (we can always take say C_1 of size δ_1^{-1}). This is a contradiction, i.e., the point x_2^1 where one has $|\widetilde{\Phi}'|\ge C_1 2^{-kl}$ for a too-large C_1 cannot exist within the integration domain.

Now in the case that $|\widetilde{\Phi}'|\gtrsim 2^{-kl}$ we may apply integration by parts and get an estimate summable in l. Let us therefore assume $|\widetilde{\Phi}'|\lesssim 2^{-kl}$, in which case we have $|\widetilde{\Phi}''|\sim 2^{-kl}$ by (4-20). Then the van der Corput lemma implies that

$$2^{-l(1/2-\sigma)(k-2)}|\hat{\nu}_l(\xi)|\lesssim(1+|\xi_3|)^{-1}.$$

The problem is now that a priori we may not sum this estimate in l. Luckily, it turns out that one can pin down the size of 2^{-l}, which in turn will pin down the number l to a finite set of size $\mathcal{O}(1)$. In order to prove this we use the expression (4-18) and the normal form of ϕ_l,

$$\phi_l(x,\delta,2^{-l})=x_1^2\tilde{r}_1(\delta_1x_1,\varepsilon x_2)+2^{-kl}x_2^k\tilde{r}_2(\delta_1x_1,\varepsilon x_2),$$

from which one has

$$(\partial_2\phi_l)(x,\delta,2^{-l})=\varepsilon x_1^2(\partial_2\tilde{r}_1)(\delta_1x_1,\varepsilon x_2)+2^{-kl}x_2^{k-1}\tilde{r}_3(\delta_1x_1,\varepsilon x_2), \tag{4-21}$$

where $\tilde{r}_3(0)\neq 0$ is a smooth function.

The idea is as follows. First, by compactness we may assume that we integrate in x_2 over a sufficiently small neighborhood of a point x_2^0 satisfying $|x_2^0|\sim 1$. In particular, we may write

$$\begin{aligned}\widetilde{\Phi}'(x_2,s_1,s_2,\delta_1,\varepsilon,2^{-kl})&=\widetilde{\Phi}'(x_2^0,s_1,s_2,\delta_1,\varepsilon,2^{-kl})+\mathcal{O}(|\widetilde{\Phi}''|)\\&=\widetilde{\Phi}'(x_2^0,s_1,s_2,\delta_1,\varepsilon,2^{-kl})+\mathcal{O}(2^{-kl}).\end{aligned}$$

Thus, it suffices to prove that

$$|\widetilde{\Phi}'(x_2^0,s_1,s_2,\delta_1,\varepsilon,2^{-kl})|=|s_2\varepsilon+\partial_2\phi_l(x_1^c,x_2^0,s_1,s_2,\delta_1,\varepsilon,2^{-kl})|\ll 2^{-kl}$$

can happen only for finitely many l. If the above inequality does not hold, then we may simply integrate by parts and are able to simply sum in l afterwards.

If we now develop both terms in $\partial_2\phi_l$ in the ε and 2^{-kl} variables (recall that x_1^c depends on both ε and 2^{-kl}), then one gets that the expression for $\widetilde{\Phi}'$ is of the form

$$s_2\varepsilon + \sum_{i=1}^{k-1} \varepsilon^i f_i(x_2^0, s_1, s_2, \delta_1) + 2^{-kl} g_0(x_2^0, s_1, s_2, \delta_1) + \mathcal{O}(2^{-kl}),$$

where we used the fact that $\varepsilon^k = (\delta_2 2^{-l})^k \ll 2^{-kl}$. Note that we have $|g_0| \sim 1$ by (4-21) (and also $|f_1| \sim 1$, but this is not important). We have to find out how many l's satisfy

$$\left| \tilde{f}_1(x_2^0, s_1, s_2, \delta_1) + \sum_{i=2}^{k-1} \varepsilon^{i-1} f_i(x_2^0, s_1, s_2, \delta_1) + \varepsilon^{-1}2^{-kl} g_0(x_2^0, s_1, s_2, \delta_1) + \mathcal{O}(\varepsilon^{-1}2^{-kl}) \right| \ll \varepsilon^{-1}2^{-kl},$$

where $\tilde{f}_1(x_2^0, s_1, s_2, \delta_1) := s_2 + f_1(x_2^0, s_1, s_2, \delta_1)$. But now one easily shows that this inequality is possible only if at least two of the terms are comparable in size (precisely because $|g_0| \sim 1$). This implies in particular that we can determine l in terms of $(x_2^0, s_1, s_2, \delta_1)$, which finishes the proof.

We mention that, interestingly, one can prove $f_2(x_2^0, s_1, s_2, 0) = 0$, a consequence of the relation (4-14).

4F. *Normal form* (vi). Here we obtain

$$\tilde{\phi}(x,\delta) = (x_2 - x_1^2\psi(\delta_1 x_1))^k r(\delta_1 x_1, \delta_2 x_2),$$
$$\mathcal{H}_{\tilde{\phi}}(x,\delta) = (x_2 - x_1^2\psi(\delta_1 x_1))^{2k-3} r_0(\delta_1 x_1, \delta_2 x_2),$$

where $r(0), r_0(0), \psi(0) \neq 0$. Thus, we may localize to the part where $|x_2 - x_1^2\psi(\delta_1 x_1)| \ll 1$; i.e., it is sufficient to consider the measure

$$f \mapsto \int f(x, \tilde{\phi}(x,\delta))\, |(x_2 - x_1^2\psi(\delta_1 x_1))^{2k-3} r_0(\delta_1 x_1, \delta_2 x_2)|^\sigma\, \chi_0(\tilde{\phi}(x,\delta))\, a(x,\delta)\, dx$$

since $|\tilde{\phi}(x,\delta)| \sim |x_2 - x_1^2\psi(\delta_1 x_1)|^k$. Note that here we have $|x_1| \sim 1 \sim |x_2|$.

Now, the next idea is to use, as in [Ikromov and Müller 2016], a Littlewood–Paley decomposition in the x_3-direction (for the mixed-norm Littlewood–Paley theory see [Lizorkin 1970]) and reduce ourselves to proving the Fourier restriction estimate for the measure piece

$$\langle \nu_l, f\rangle = \int f(x, \tilde{\phi}(x,\delta))\, |(x_2 - x_1^2\psi(\delta_1 x_1))^{2k-3} r_0(\delta_1 x_1, \delta_2 x_2)|^\sigma\, \chi_1(2^{kl}(\tilde{\phi}(x,\delta)))\, a(x,\delta)\, dx.$$

Using the coordinate transformation $x_2 \mapsto x_2 + x_1^2\psi(\delta_1 x_1)$ we may write

$$\begin{aligned}\langle \nu_l, f\rangle = \int f(x_1, x_2 + x_1^2\psi(\delta_1 x_1), x_2^k \tilde{r}(\delta_1 x_1, \delta_2 x_2)) \\ \times |x_2^{2k-3}\tilde{r}_0(\delta_1 x_1, \delta_2 x_2)|^\sigma\, \chi_1(2^{kl} x_2^k \tilde{r}(\delta_1 x_1, \delta_2 x_2))\, \tilde{a}(x,\delta)\, dx,\end{aligned}$$

where $|\tilde{r}| \sim 1$ is a smooth function. Finally, we use the coordinate transformation $x_2 \mapsto 2^{-l}x_2$ and rescale f in the third coordinate. Then we are reduced to proving the Fourier restriction estimate

$$\|\hat{f}\|^2_{L^2(d\tilde{\nu}_{j,l})} \le C 2^{l(1-3\sigma)} \|f\|^2_{L^{\mathfrak{p}_3}_{x_3}(L^{\mathfrak{p}_1}_{(x_1,x_2)})} \tag{4-22}$$

for the measure

$$\langle \tilde{\nu}_{j,l}, f \rangle = \int f(x_1, 2^{-l}x_2 + x_1^2\psi(\delta_1 x_1), x_2^k \tilde{r}(\delta_1 x_1, 2^{-l}\delta_2 x_2))\, a(x, \delta, 2^{-l})\, \mathrm{d}x, \tag{4-23}$$

where a is supported so that $|x_1| \sim 1$ and $|x_2| \sim 1$. Now we note that the estimate for $\sigma = 0$ follows by Plancherel, while the estimate for $\sigma = \frac{1}{3}$ is going to be shown in Section 5 since the form of the measure $\tilde{\nu}_{j,l}$ coincides with the form in (5-11) below. Interpolating, we obtain the estimate for all $0 \le \sigma \le \frac{1}{3}$.

Note that when

$$\frac{1}{p_1'} = \frac{1}{p_3'} = \frac{1}{4},$$

one can simplify the proof by a modification of Lemma 4.2, i.e., by using the Fourier decay of $\tilde{\nu}_{j,l}$, which is easily seen to be

$$|\hat{\tilde{\nu}}_{j,l}(\xi)| \lesssim 2^{l/2}(1 + |\xi|)^{-1},$$

and by using the Plancherel theorem, but this time in the (x_1, x_3)-plane (which is why it works only for $1/p_1' = 1/p_3'$) since the mapping $(x_1, x_2) \mapsto (x_1, x_2^k \tilde{r}(\delta_1 x_1, 2^{-l}\delta_2 x_2))$ has Jacobian of size ~ 1. In fact, in Section 5 we shall combine this idea of using Lemma 4.2 with the methods used in [Ikromov and Müller 2016] (and [Palle 2021]).

4F1. *A Knapp-type example.* Let us now show by using a Knapp-type example that one cannot get the estimate

$$\|\hat{f}\|_{L^2(\mathrm{d}\nu)} \le C \|f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})}$$

for $\sigma > \frac{1}{3}$ where ν is the surface measure

$$\langle \nu, f \rangle = \int_{\mathbb{R}^2 \setminus \{0\}} f(x, \phi_{\mathrm{loc}}(x))\, a(x)\, |\mathcal{H}_\phi(x)|^\sigma\, \mathrm{d}x$$

and ϕ_{loc} is given by the normal form (vi). Let us consider the function $f = \varphi_\varepsilon$ defined by

$$\hat{\varphi}_\varepsilon(x) = \chi_0\Big(\frac{x_1}{\varepsilon^\delta}\Big)\chi_0\Big(\frac{x_2}{\varepsilon^{2\delta}}\Big)\chi_0\Big(\frac{x_3}{\varepsilon}\Big)$$

for some small ε and δ. Its mixed L^p norm is

$$\|\varphi_\varepsilon\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})} \sim \varepsilon^{3\delta/p_1' + 1/p_3'}.$$

Now, in the integral

$$\int |\hat{\varphi}_\varepsilon|^2\, \mathrm{d}\nu = \int |\hat{\varphi}_\varepsilon|^2(x, \phi_{\mathrm{loc}}(x)) a(x) |\mathcal{H}_{\phi_{\mathrm{loc}}}(x)|^\sigma\, \mathrm{d}x$$

we integrate over the set

$$D_\varepsilon^0 := \{x \in \mathbb{R}^2 : |x_1| \lesssim \varepsilon^\delta, \quad |x_2| \lesssim \varepsilon^{2\delta}, \quad |\phi_{\mathrm{loc}}(x)| \sim |x_2 - x_1^2\psi(x_1)|^k \lesssim \varepsilon\}$$

by the definition of φ_ε. If δ is sufficiently small, D_ε^0 contains the set

$$D_\varepsilon := \{x \in \mathbb{R}^2 : |x_1| \lesssim \varepsilon^\delta, \quad |\phi_{\mathrm{loc}}(x)| \sim |x_2 - x_1^2\psi(x_1)| \lesssim \varepsilon^{1/k}\},$$

and so if the Fourier restriction estimate holds, one has

$$\varepsilon^{6\delta/p_1'+2/p_3'} \sim \|\varphi_\varepsilon\|^2_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})} \gtrsim \int |\hat{\varphi}_\varepsilon|^2 \,\mathrm{d}\nu \gtrsim \int_{D_\varepsilon} |x_2 - x_1^2\psi(x_1)|^{\sigma(2k-3)} \,\mathrm{d}x$$
$$\sim \varepsilon^\delta \int_{|y|\lesssim \varepsilon^{1/k}} |y|^{\sigma(2k-3)} \,\mathrm{d}y \sim \varepsilon^{\delta+(\sigma(2k-3)+1)/k}.$$

Letting ε and then δ tend to 0 we obtain the condition

$$\frac{1}{p_3'} \le \frac{\sigma(2k-3)+1}{2k} = \sigma + \frac{1-3\sigma}{2k}.$$

Since we are interested in

$$\frac{1}{p_1'} = \frac{1}{2} - \sigma, \qquad \frac{1}{p_3'} = \sigma,$$

the above inequality implies precisely $\sigma \le \frac{1}{3}$.

5. Fourier restriction without a mitigating factor

Here we prove Theorem 1.2, i.e., the estimate

$$\|\hat{f}\|_{L^2(\mathrm{d}\mu)} \le C\|f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})}$$

for μ the surface measure of the form

$$\langle \mu, f\rangle = \int_{\mathbb{R}^2\setminus\{0\}} f(x,\phi(x))\, |x|_\alpha^{2\vartheta} \,\mathrm{d}x,$$

where

$$\vartheta = \frac{|\alpha|}{p_1'} + \frac{\rho}{p_3'} - \frac{|\alpha|}{2}.$$

Recall that this ϑ is chosen (depending on $(p_1, p_3) \in (1,2]^2$) precisely so that the above restriction estimate is equivalent to the local estimate

$$\|\hat{f}\|_{L^2(\mathrm{d}\mu_0)} \le C\|f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})},$$

where μ_0 is the surface measure

$$\langle \mu_0, f\rangle = \int_{\mathbb{R}^2\setminus\{0\}} f(x,\phi(x))\, \eta(x)\, |x|_\alpha^{2\vartheta} \,\mathrm{d}x \tag{5-1}$$

for $\eta \in C_c^\infty(\mathbb{R}^2 \setminus \{0\})$ identically equal to 1 in an annulus.

Note that $|x|_\alpha^{2\vartheta}$ is not smooth near the axes. Luckily, we shall be able to circumvent this problem by using the Littlewood–Paley theorem to localize away from the axes, as was done in the case with the mitigating factor.

Now we recall the necessary conditions from [Palle 2021, Proposition 2.1] obtained through the Knapp-type examples. Let us fix a point v such that $\eta(v) \neq 0$ and let η_v be a smooth cutoff function identically equal to η on a small neighborhood of v. It suffices to consider the measure

$$\langle \mu_{0,v}, f\rangle = \int_{\mathbb{R}^2\setminus\{0\}} f(x, \phi_v(x-v))\, \eta_v(x)\, |x|_\alpha^{2\vartheta}\, \mathrm{d}x, \tag{5-2}$$

where we recall from the Introduction that

$$\phi_v(x) = \phi(x+v) - \phi(v) - x\cdot\nabla\phi(v).$$

We recall also that $h_{\mathrm{lin}}(\phi, v)$ is the linear height of ϕ_v at its origin, and that $h(\phi, v)$ is its Newton height.

Proposition 2.1 of [Palle 2021] tells us what the necessary conditions on the exponents p_1 and p_3 are if the $L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)}) \to L^2(\mathrm{d}\mu_{0,v})$ Fourier restriction estimate were to hold true. The input data one needs is the Newton polyhedron of the phase function ϕ_v at the origin in both its linearly adapted and adapted coordinates. When the linearly adapted and adapted coordinates do not coincide, one constructs from the two Newton polyhedra the so-called augmented Newton polyhedron. When the linearly adapted and adapted coordinates do coincide, then one obtains a single condition, namely,

$$\frac{1}{p_1'} + \frac{h_{\mathrm{lin}}(\phi, v)}{p_3'} \le \frac{1}{2}. \tag{5-3}$$

Otherwise, in the proposition it is shown that to each edge of the augmented Newton polyhedron, say contained in the line $\{(t_1, t_2) \in \mathbb{R}^2 : \tilde{\kappa}_1 t_1 + \tilde{\kappa}_2 t_2 = 1\}$, one can associate the necessary condition

$$\frac{(1+m)\tilde{\kappa}_1}{p_1'} + \frac{1}{p_3'} \le \frac{\tilde{\kappa}_1 + \tilde{\kappa}_2}{2},$$

where m is the negative reciprocal of the slope of the principal face of the Newton polyhedron of ϕ_v in its linearly adapted coordinates. As shown in [Palle 2021, Proposition 2.1], this set of conditions always contains the condition (5-3) and the condition

$$\frac{1}{p_3'} \le \frac{1}{2h(\phi, v)}. \tag{5-4}$$

Thus, if ϕ satisfies (LA) at v, then $h_{\mathrm{lin}}(\phi, v) = h(\phi, v)$, and the only necessary condition is given by (5-3). If ϕ does not satisfy (LA) at v, then from Proposition 1.4 we deduce that out of all the normal forms this is only possible for the normal form

$$\phi_{v,y}(y) := (y_2 - y_1^2\psi(y_1))^k r(y),$$

where $r(0) \neq 0$, $\psi(0) \neq 0$, and $2 \le k < \infty$, since all the normal forms are linearly adapted and this is the only nonadapted normal form (see [Ikromov and Müller 2011], or the Introduction of [Ikromov and Müller 2016] to find precise conditions for whether a function is in linearly adapted or adapted coordinates). Using this normal form one can now determine its augmented Newton polyhedron, which turns out to have only two edges. Its two associated conditions turn out to be precisely the conditions (5-3) and (5-4), i.e.,

$$\frac{1}{p_1'} + \frac{h_{\mathrm{lin}}(\phi, v)}{p_3'} \le \frac{1}{2} \quad \text{and} \quad \frac{h(\phi, v)}{p_3'} \le \frac{1}{2}. \tag{5-5}$$

One also easily shows $h(\phi, v) = k$ and $h_{\text{lin}}(\phi, v) = 2k/3$. Note that in the case $h_{\text{lin}}(\phi, v) = h(\phi, v)$ the second condition in (5-5) would be redundant. Thus, if we now vary v over the points where $\eta(v) \neq 0$, then we obtain the conditions

$$\frac{1}{p_1'} + \frac{h_{\text{lin}}(\phi)}{p_3'} \leq \frac{1}{2} \quad \text{and} \quad \frac{h(\phi)}{p_3'} \leq \frac{1}{2},$$

where we remind that $h_{\text{lin}}(\phi)$ and $h(\phi)$ are respectively global linear height and global Newton height defined as in (1-6).

At all points v where (LA) is satisfied and where $|x|_\alpha^{2\vartheta}$ is smooth (i.e., v is not on an axis) we get the local Fourier restriction estimate in the range (5-3) directly from [Palle 2021, Proposition 4.2]. We shall briefly touch upon what happens in the case when v is situated on the axis in Section 5A. In this case one has to only slightly adjust the proofs in Section 4.

In the case when (LA) is not satisfied at v let us call the pair $(\mathfrak{p}_1, \mathfrak{p}_3) = (\mathfrak{p}_1(v), \mathfrak{p}_3(v))$ given by

$$\left(\frac{1}{\mathfrak{p}_1'}, \frac{1}{\mathfrak{p}_3'}\right) = \left(\frac{1}{2} - \frac{h_{\text{lin}}(\phi, v)}{2h(\phi, v)}, \frac{1}{2h(\phi, v)}\right)$$

the critical exponent of ϕ at v. It is obtained as the intersection of the lines

$$\frac{1}{p_1'} + \frac{h_{\text{lin}}(\phi, v)}{p_3'} = \frac{1}{2} \quad \text{and} \quad \frac{1}{p_3'} = \frac{1}{2h(\phi, v)}$$

in the $(1/p_1', 1/p_3')$ plane. Thus, for the local estimate in this case it suffices to prove the inequality

$$\|\hat{f}\|_{L^2(\mathrm{d}\mu_{0,v})} \leq C \|f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})},$$

where

$$\langle \mu_{0,v}, f \rangle = \int f(x_1, x_2, \phi_v(x_1, x_2))\, \eta_v(x_1, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2$$

and

$$\left(\frac{1}{p_1'}, 1p_3'\right) \in \left\{\left(0, \frac{1}{2h(\phi, v)}\right), \left(\frac{1}{2}, 0'\right), \left(\frac{1}{\mathfrak{p}_1'(v)}, \frac{1}{\mathfrak{p}_3'(v)}\right)\right\},$$

since then we get the full range from the necessary conditions by interpolation. We shall only give a sketch of the proof in this case too in Subsections 5B and 5C, since it is almost identical to a type of singularity considered in [Palle 2021, Section 5.5].

5A. ***Fourier restriction for the adapted case.*** As mentioned, in the adapted case one needs to prove the Fourier restriction estimate for $(p_1, p_3) \in (1, 2)^2$ satisfying

$$\frac{1}{p_1'} + \frac{h_{\text{lin}}(\phi, v)}{p_3'} = \frac{1}{2},$$

and the part of the measure where the amplitude in (5-1) is smooth the restriction estimate is already proven in [Palle 2021].

Now the amplitude in (5-1) (in particular the function $x \mapsto |x|_\alpha^{2\vartheta}$) is in general not smooth along the axes $x_1 = 0$ and $x_2 = 0$. Namely, on the $x_1 = 0$ axis one can take only the derivatives (of the amplitude) in the x_2-direction, and analogously on the $x_2 = 0$ axis one can take only derivatives in the x_1-direction. Note that the only possible nonadapted normal form appears only away from the axes.

Let us consider without loss of generality what happens for the point $v = (v_1, 0)$ on the axis $x_2 = 0$ and its associated measure $\mu_{0,v}$ defined in (5-2). We shall only briefly sketch what one needs to do in order to prove the Fourier restriction estimate when the amplitude is not smooth in the x_2-direction at v. Since we are dealing only with adapted normal forms, it suffices to obtain an appropriate estimate on the Fourier transform, after which one can apply Lemma 4.2 or its modification such as [Palle 2021, Lemma 3.8]. For the reader's convenience we state explicitly the result we need (the proof is essentially the same as for Lemma 4.2 — in fact it is even simpler since one can use the usual Young's inequality instead of the weak one).

Lemma 5.1. *Assume that we are given a bounded open set $\Omega \subseteq \mathbb{R}^2$ and functions $\Phi \in C^\infty(\Omega; \mathbb{R}^2)$, $\phi \in C^\infty(\Omega; \mathbb{R})$, $a \in L^\infty(\Omega)$. Let us consider the measure*

$$\langle \mu, f \rangle := \int f(\Phi(x), \phi(x))\, a(x)\, \mathrm{d}x$$

and the operator $T : f \mapsto f * \hat{\mu}$. If Φ is injective, its Jacobian is of size $|J_\Phi| \sim A_1$, and if one has furthermore the estimate

$$|\hat{\mu}(\xi)| \leq A_2(1 + |\xi_3|)^{-1/h}$$

for some $h \in (0, 1)$, then for any $\theta \in [0, 1]$ and

$$\left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1-\theta}{2}, \frac{\theta}{2h}\right)$$

the $L^{p_3}_{x_3}(\mathbb{R}; L^{p_1}_{(x_1,x_2)}(\mathbb{R}^2)) \to L^{p_3'}_{x_3}(\mathbb{R}; L^{p_1'}_{(x_1,x_2)}(\mathbb{R}^2))$ operator norm of T is bounded by $(A_1^{-1}\|a\|_{L^\infty})^{1-\theta} A_2^\theta$.

Often we shall also need to use the Littlewood–Paley theorem in order to localize away from the axis.

According to the normal forms listed at the end of Section 3A, and under the condition (H1), we have the following cases.

Case 1: If (under the notation of Section 3) we have $k = \infty$, then by the considerations from Section 3B the phase at v is

$$\phi_v(x - v) = (x_1 - v_1)^{\tilde{k}} q(x_1 - v_1) + \varphi(x_1, x_2),$$

where $2 \leq \tilde{k} < \infty$, $q(0) \neq 0$, and φ is a flat function at v. This corresponds to normal form (i.y2) and we have $h_{\mathrm{lin}}(\phi, v) = \tilde{k}$. Since $|x|_\alpha^{2\vartheta}$ is still smooth in the x_1-direction, one can use the van der Corput lemma in the x_1-direction and get that the decay of the Fourier transform of $\mu_{0,v}$ is $(1 + |\xi|)^{-1/\tilde{k}}$. This now implies the desired estimate by Lemma 5.1.

If $2 \leq k < \infty$, then we have three further cases.

Case 2: Let us consider the phase

$$\phi_v(x) = x_2^k r(x),$$

where $r(v) \neq 0$ and $k \geq 2$. In this case the linear height is $h_{\text{lin}}(\phi, v) = k$. Here the idea is to apply the Littlewood–Paley theorem in order to localize away from the axis $x_2 = 0$, and rescale afterwards. Since essentially the same thing was done in Section 4 for this type of singularity (see the proof for normal form (i) in Section 4A), let us just briefly mention the main differences compared to there. Obviously, one scales differently the measure pieces away from the axis obtained by applying the Littlewood–Paley theorem since here we consider different exponents (p_1, p_3). The main difference is that we do not use the Hessian determinant to obtain a decay on the Fourier transformation of the rescaled measure piece (since the Hessian determinant may vanish of infinite order as only (H1) is assumed and not the stronger condition (H2)), but rather determine it directly from the form of the phase above. This we may now do since the new amplitude for the rescaled measure pieces is now smooth.

Case 3: Let us now consider the case when the phase is nondegenerate, i.e., the Hessian determinant does not vanish at v (and in particular $h_{\text{lin}}(\phi, v) = 1$). Here we use the Littlewood–Paley theorem as in Case 2, but after rescaling we use the size of the Hessian determinant of the new phase to get a decay on the Fourier transform of the measure (as was done in Section 4 for normal forms (i), (ii), and (iii)).

Case 4: The final case is when (after an affine change to y- or w-coordinates from Section 3) we have

$$\phi_{v,u}(u) = u_1^2 r_1(u) + u_2^{k_0} r_2(u),$$

where $3 \leq k_0 \leq \infty$, $r_1(0) \neq 0$, and in the case when $k_0 < \infty$ we have $r_2(0) \neq 0$ and $h_{\text{lin}}(\phi, v) = 2k_0/(2+k_0)$. If $k_0 = \infty$ then $h_{\text{lin}}(\phi, v) = 2$, and the above equality holds in the sense that we can take any $k_0 \geq 0$ and r_2 flat at the origin. Inspecting the y- and w-coordinates from Section 3 we see that the $x_2 = 0$ axis corresponds to the $u_2 = 0$ axis.

If $k_0 = \infty$, we can argue in the same way as in the case $k = \infty$ above (here it is critical that $\partial_{u_1} = c\partial_{x_1}$, $c \neq 0$, in order to be able to apply the van der Corput lemma in the smooth direction).

Otherwise, if k_0 is finite, we proceed again with a Littlewood–Paley decomposition in the u_2-direction (as was done in Section 4C for normal forms (ii) and (iii)) in order to get a smooth amplitude. At this point one gets that the estimate on the decay of the Fourier transform is $2^{k_0 l/2}(1 + |\xi|)^{-1}$ by using the size of the Hessian determinant. Since the new rescaled phase is (compare with (4-9))

$$u_1^2 r_1(u_1, 2^{-l} u_2) + 2^{-k_0 l} u_2^{k_0} r_2(u_1, 2^{-l} u_2),$$

by applying the van der Corput lemma in u_1 we also have the decay estimate $(1 + |\xi|)^{-1/2}$. Interpolating these two estimates gives the decay $2^l (1 + |\xi|)^{-(2+k_0)/(2k_0)}$, which turns out to be precisely what one needs when applying Lemma 5.1.

5B. *Fourier restriction for the nonadapted case: preliminaries.* Let us fix a phase function ϕ_{loc} of the form

$$\phi_{\text{loc}}(x) = (x_2 - x_1^2 \psi(x_1))^k r(x),$$

where $\psi(0), r(0) \neq 0$ and $k \in \mathbb{N}$, $k \geq 2$. The adapted coordinates are obtained by the smooth transformation $y_1 = x_1$, $y_2 = x_2 - x_1^2 \psi(x_1)$:

$$\phi_{\text{loc}}^a(y) := y_2^k r^a(y),$$

where $r^a(0) \neq 0$. Thus, the Newton height of ϕ_{loc} is k and the Newton distance is $d := 2k/3$ (which coincides with the linear height h_{lin}). The Varchenko exponent is 0 since in adapted coordinates the principal face is noncompact. Then from, e.g., [Palle 2021, Section 3.3] we know that we automatically have the Fourier restriction estimate

$$\|\mathcal{F} f\|_{L^2(\mathrm{d}\nu)} \lesssim \|f\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)})}, \quad f \in \mathcal{S}(\mathbb{R}^3), \tag{5-6}$$

for the exponents

$$\left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(0, \frac{1}{2k}\right) \quad \text{and} \quad \left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1}{2}, 0\right),$$

and where the measure ν is defined through

$$\langle \nu, f\rangle = \int f(x_1, x_2, \phi_{\mathrm{loc}}(x_1, x_2))\, a(x_1, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2, \tag{5-7}$$

where $a \in C_c^\infty(\mathbb{R}^2)$ is a nonnegative function supported in a small neighborhood of the origin. It remains to obtain the Fourier restriction estimate for the critical exponent, which in this case is

$$\left(\frac{1}{\mathfrak{p}_1'}, \frac{1}{\mathfrak{p}_3'}\right) = \left(\frac{1}{6}, \frac{1}{2k}\right). \tag{5-8}$$

The case $k = 2$ has been solved in [Palle 2021]. In the case $k = 3$ the critical exponent lies on the diagonal and so this case has already been solved in [Ikromov and Müller 2016].

In the case $k \geq 4$ we have $1/\mathfrak{p}_1' > 1/\mathfrak{p}_3'$ and so one would need to slightly modify the methods used in [Palle 2021] (i.e., the methods for the case $h_{\mathrm{lin}}(\phi) < 2$) since there one interpolated between two points of the form

$$\left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(0, \frac{s}{2}\right) \quad \text{and} \quad \left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1}{2}, \frac{1}{2}\right)$$

for some $0 < s < 1/k$. In the case $1/\mathfrak{p}_1' > 1/\mathfrak{p}_3'$ in general one would need to interpolate between three points

$$\left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = (0,0), \quad \left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1}{2}, \frac{1}{2}\right), \quad \text{and} \quad \left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1}{2}, 0\right).$$

In particular, if one has an operator $T : L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)}) \to L^{p_3'}_{x_3}(L^{p_1'}_{(x_1,x_2)})$ satisfying the estimates

$$\begin{aligned}
\|T\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)}) \to L^{p_3'}_{x_3}(L^{p_1'}_{(x_1,x_2)})} &\lesssim A_1 \quad \text{for } \left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = (0,0),\\
\|T\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)}) \to L^{p_3'}_{x_3}(L^{p_1'}_{(x_1,x_2)})} &\lesssim A_2 \quad \text{for } \left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1}{2}, \frac{1}{2}\right),\\
\|T\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)}) \to L^{p_3'}_{x_3}(L^{p_1'}_{(x_1,x_2)})} &\lesssim A_3 \quad \text{for } \left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1}{2}, 0\right),
\end{aligned} \tag{5-9}$$

then by interpolation one has the estimate

$$\|T\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)}) \to L^{p_3'}_{x_3}(L^{p_1'}_{(x_1,x_2)})} \lesssim A_1^{2/3}\, A_2^{1/k}\, A_3^{(k-3)/(3k)} \quad \text{for } \left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1}{6}, \frac{1}{2k}\right).$$

In our special case we shall not use the above general approach since we recall that when we considered the case when the mitigating factor was present (to be more precise, the case of normal form (vi) considered in Section 4F), after performing some decompositions and rescalings one got measure pieces for which one needed to prove the Fourier restriction estimate for the exponent

$$\left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1}{6}, \frac{1}{3}\right). \tag{5-10}$$

In the current case without the mitigating factor it turns out that we shall get the same measure pieces, but for which we need to prove the Fourier restriction estimate for the exponent (5-8). Thus, if we have the Fourier restriction estimate for the exponent (5-10), then the Fourier restriction for (5-8) is obtained by interpolating with the result for

$$\left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1}{6}, 0\right),$$

which one can obtain by applying the 2-dimensional Fourier restriction result for curves with nonvanishing curvature.

These stronger estimates for the rescaled measure pieces do not contradict the necessary conditions obtained by Knapp-type examples in [Palle 2021] since the information on the exponents and the Newton height of ϕ is consumed in the rescaling procedure (which is different in this section and in Section 4F).

Let us begin with some preliminary reductions. By the results from [Palle 2021, Section 4.2], instead of considering the whole measure (5-7), we may reduce ourselves to considering the part near the principal root jet in the half-plane $\{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0\}$:

$$\langle \nu^{\rho_1}, f \rangle = \int_{x_1 \geq 0} f(x, \phi_{\mathrm{loc}}(x))\, a(x)\, \rho_1(x)\, \mathrm{d}x,$$

where

$$\rho_1(x) = \chi_0\left(\frac{x_2 - \psi(0)x_1^2}{\varepsilon x_1^2}\right)$$

for an ε which we can take to be as small as we want.

The next step is to use a Littlewood–Paley argument in the (x_1, x_2)-plane and the scaling by κ dilations

$$\delta_t^\kappa(x) = (t^{\kappa_1} x_1, t^{\kappa_2} x_2),$$

where $\kappa := (1/(2k), 1/k)$ is the weight associated to the principal face of ϕ_{loc}. Then one is reduced to proving (5-6) for the measures

$$\langle \nu_j, f \rangle = \int f(x, \phi(x, \delta))\, a(x, \delta)\, \mathrm{d}x,$$

uniformly in j, where the function $\phi(x, \delta)$ has the form

$$\phi(x, \delta) := (x_2 - x_1^2 \psi(\delta_1 x_1))^k r(\delta_1 x_1, \delta_2 x_2),$$

where

$$\delta = (\delta_1, \delta_2) := (2^{-\kappa_1 j}, 2^{-\kappa_2 j}).$$

Note that we can take $|\delta| \ll 1$. The amplitude $a(x,\delta) \geq 0$ is a smooth function of (x,δ) supported where

$$x_1 \sim 1 \sim |x_2|.$$

We may additionally assume $|x_2 - x_1^2\psi(0)| \ll 1$ due to ρ_1, and by compactness we may in fact reduce ourselves to assuming $|(x_1,x_2)-(v_1^0,v_2^0)| \ll 1$ for some $(v_1^0,v_2^0) \in \mathbb{R}^2$ with $v_1^0 \sim 1$.

The following step is to again apply the Littlewood–Paley theorem, but this time in the x_3-direction (again, for the mixed-norm Littlewood–Paley theory see [Lizorkin 1970]), and reduce the Fourier restriction problem for ν_j to the Fourier restriction for the measures

$$\langle \nu_{\delta,l}, f\rangle = \int f(x,\phi(x,\delta))\,\chi_1(2^{kl}\phi(x,\delta))\,a(x,\delta)\,\mathrm{d}x;$$

i.e., we need to prove

$$\|\mathcal{F}f\|_{L^2(\mathrm{d}\nu_{\delta,l})} \lesssim \|f\|_{L^{\mathfrak{p}_3}_{x_3}(L^{\mathfrak{p}_1}_{(x_1,x_2)})}, \quad f \in \mathcal{S}(\mathbb{R}^3),$$

uniformly in l and δ, where $l \gg 1$ and $|\delta| \ll 1$.

Finally, we perform a change of coordinates and a rescaling. Namely, after substituting $(x_1,x_2) \mapsto (x_1, 2^{-l}x_2 + x_1^2\psi(\delta_1 x_1))$ we get

$$\langle \nu_{\delta,l}, f\rangle = 2^{-l}\int f(x_1, 2^{-l}x_2 + x_1^2\psi(\delta_1x_1), 2^{-kl}\phi^a(x,\delta,l))\,a(x,\delta,l)\,\mathrm{d}x,$$

where

$$\begin{aligned} a(x,\delta,l) &:= \chi_1(\phi^a(x,\delta,l))\,a(x_1, 2^{-l}x_2 + x_1^2\psi(\delta_1x_1),\delta),\\ \phi^a(x,\delta,l) &:= x_2^k r(\delta_1x_1, \delta_2(2^{-l}x_2 + x_1^2\psi(\delta_1x_1))). \end{aligned}$$

Note that $a(x,\delta,l)$ is again supported in a domain where $x_1 \sim 1 \sim |x_2|$. Rescaling we obtain that the Fourier restriction estimate for $\nu_{\delta,l}$ is equivalent to the estimate

$$\|\mathcal{F}f\|_{L^2(\mathrm{d}\tilde{\nu}_{\delta,l})} \lesssim \|f\|_{L^{\mathfrak{p}_3}_{x_3}(L^{\mathfrak{p}_1}_{(x_1,x_2)})}, \quad f \in \mathcal{S}(\mathbb{R}^3),$$

for the measure

$$\langle \tilde{\nu}_{\delta,l}, f\rangle = \int f(x_1, 2^{-l}x_2 + x_1^2\psi(\delta_1x_1), \phi^a(x,\delta,l))\,a(x,\delta,l)\,\mathrm{d}x. \tag{5-11}$$

As mentioned, since this measure is of the same form as (4-23), we are interested in proving the stronger estimate

$$\|\mathcal{F}f\|_{L^2(\mathrm{d}\tilde{\nu}_{\delta,l})} \lesssim \|f\|_{L^{\tilde{\mathfrak{p}}_3}_{x_3}(L^{\tilde{\mathfrak{p}}_1}_{(x_1,x_2)})}, \quad f \in \mathcal{S}(\mathbb{R}^3),$$

where

$$\left(\frac{1}{\tilde{\mathfrak{p}}_1'}, \frac{1}{\tilde{\mathfrak{p}}_3'}\right) := \left(\frac{1}{6}, \frac{1}{3}\right).$$

Note that we automatically have the estimate for

$$\left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1}{6}, 0\right)$$

by a classical result of Fefferman and Stein [Fefferman 1970] (or [Zygmund 1974]), since $x_1 \mapsto (x_1, 2^{-l}x_2 + x_1^2\psi(\delta_1 x_1))$ is a curve with curvature bounded from below uniformly in $|x_2| \sim 1$, $2^{-l} \ll 1$, and $\delta_1 \ll 1$.

5C. ***Fourier restriction for the nonadapted case*: *spectral decomposition.*** We begin by performing a spectral decomposition of the measure $\tilde{\nu}_{\delta,l}$. For $(\lambda_1, \lambda_2, \lambda_3)$ dyadic numbers with $\lambda_i \geq 1$, $i = 1, 2, 3$, we consider localized measures ν_l^λ defined through

$$\hat{\nu}_l^\lambda(\xi) = \chi_1\left(\frac{\xi_1}{\lambda_1}\right)\chi_1\left(\frac{\xi_2}{\lambda_2}\right)\chi_1\left(\frac{\xi_3}{\lambda_3}\right)\int e^{-i\Phi(x,\delta,l,\xi)}\, a(x,\delta,l)\,\chi_1(x_1)\,\chi_1(x_2)\,\mathrm{d}x, \tag{5-12}$$

where the phase function is

$$\Phi(x,\delta,l,\xi) := \xi_3\phi^a(x,\delta,l) + 2^{-l}\xi_2 x_2 + \xi_2 x_1^2\psi(\delta_1 x_1) + \xi_1 x_1. \tag{5-13}$$

By an abuse of notation, above whenever $\lambda_i = 1$, we consider the cutoff function $\chi_1(\xi_i/\lambda_i)$ to be actually $\chi_0(\xi_1/\lambda_1)$; i.e., it localizes so that $|\xi_i| \lesssim 1$.

Let us introduce the convolution operators $\widetilde{T}_{\delta,l} f := f * \hat{\tilde{\nu}}_{\delta,l}$ and $T_l^\lambda f := f * \hat{\nu}_l^\lambda$. Then we need to show

$$\|\widetilde{T}_{\delta,l}\|_{L_{x_3}^{\tilde{p}_3}(L_{(x_1,x_2)}^{\tilde{p}_1}) \to L_{x_3}^{\tilde{p}_3'}(L_{(x_1,x_2)}^{\tilde{p}_1'})} \lesssim 1,$$

since $\widetilde{T}_{\delta,l}$ is the "R^*R" operator, i.e., one has $\widetilde{T}_{\delta_l} = (\widetilde{R}_{\delta,l})^*\widetilde{R}_{\delta,l}$ if $\widetilde{R}_{\delta,l}$ denotes the Fourier restriction operator with respect to the surface measure $\tilde{\nu}_{\delta,l}$. Therefore, the boundedness of $\widetilde{T}_{\delta,l}$ is equivalent to the boundedness of $\widetilde{R}_{\delta,l}$ by Hölder's inequality.

Our first step shall be to reduce the problem to the case when $\lambda_2 \ll 2^l$. In order to achieve this we split the Fourier transform of $\tilde{\nu}_{\delta,l}$ as

$$\hat{\tilde{\nu}}_{\delta,l} = (1 - \chi_0(2^{-l}\xi_2))\hat{\tilde{\nu}}_{\delta,l} + \chi_0(2^{-l}\xi_2)\hat{\tilde{\nu}}_{\delta,l}, \tag{5-14}$$

where we assume that χ_0 is supported in a sufficiently small neighborhood of the origin, and we denote the respective operators for the respective terms by T_I and T_{II}.

For the first term in (5-14) and its operator T_I one uses Lemma 4.2 above, though with a slight modification. First, since on the support of $(1 - \chi_0(2^{-l}\xi_2))\hat{\tilde{\nu}}_{\delta,l}$ we have $|\xi_2| \gtrsim 2^l$, one can easily show by using (5-13) that now

$$|(1 - \chi_0(2^{-l}\xi_2))\hat{\tilde{\nu}}_{\delta,l}| \lesssim 2^{-l/2}(1 + |\xi_3|)^{-1},$$

as the "worst case" is when $|\xi_1| \sim |\xi_2|$ and $|\xi_3| \sim |2^{-l}\xi_2|$, in which case we use stationary phase in both x_1 and x_2 (and in other cases we get a better decay by integrating by parts). In order to obtain the Plancherel estimate $L^1(\mathbb{R}; L^2(\mathbb{R}^2)) \to L^\infty(\mathbb{R}; L^2(\mathbb{R}^2))$ in Lemma 4.2 for T_I it suffices to prove it for T_{II} and $\widetilde{T}_{\delta,l}$ (formally, one needs to actually consider the $L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ estimate for a fixed ξ_3). For the operator $\widetilde{T}_{\delta,l}$ we get the bound 2^l in the same way as in Lemma 4.2. The main fact to notice is that in (5-11) the Jacobian of $(x_1, x_2) \mapsto (x_1, 2^{-l}x_2 + x_1^2\psi(\delta_1 x_1))$ is of size 2^{-l}. One now gets the same estimate automatically for T_{II} since the L^1 norm of the Fourier transform of the cutoff function

$\chi_0(2^{-l}\xi_2)$ is of size ~ 1. The $L^{\tilde{p}_3}_{x_3}(L^{\tilde{p}_1}_{(x_1,x_2)}) \to L^{\tilde{p}'_3}_{x_3}(L^{\tilde{p}'_1}_{(x_1,x_2)})$ estimate for T_I follows with constant of size $\sim 1 = (2^{-l/2})^{2/3}(2^l)^{1/3}$.

For the operator T_{II} we shall use the spectral decomposition (5-12) where we may now assume $\lambda_2 \ll 2^l$. Recall that for an operator of the form $Tf = f * \hat{g}$ the A_1-constant from (5-9) is bounded by the L^∞ norm of $\hat{g}$, and the A_2-constant is bounded by the L^∞ norm of g. If we now furthermore have that $\hat{g}$ has its support in the ξ_3-coordinate localized at $|\xi_3| \lesssim \lambda_3$, then by [Palle 2021, Lemma 3.9] we have the estimate

$$\|T\|_{L^{p_3}_{x_3}(L^{p_1}_{(x_1,x_2)}) \to L^{p'_3}_{x_3}(L^{p'_1}_{(x_1,x_2)})} \lesssim A_1 \lambda_3^{1/2} \quad \text{for } \left(\frac{1}{p'_1}, \frac{1}{p'_3}\right) = \left(0, \frac{1}{4}\right),$$

and so by interpolation we get

$$\|T\|_{L^{\tilde{p}_3}_{x_3}(L^{\tilde{p}_1}_{(x_1,x_2)}) \to L^{\tilde{p}'_3}_{x_3}(L^{\tilde{p}'_1}_{(x_1,x_2)})} \lesssim A_1^{2/3} A_2^{1/3} \lambda_3^{1/3}. \tag{5-15}$$

The inverse Fourier transform of (5-12) is

$$\nu_l^\lambda(x) = \lambda_1\lambda_2\lambda_3 \int \check{\chi}_1(\lambda_1(x_1 - y_1))\, \check{\chi}_1(\lambda_2(x_2 - 2^{-l} y_2 - y_1^2 \psi(\delta_1 y_1))) \\ \times \check{\chi}_1(\lambda_3(x_3 - \phi^a(y, \delta, l)))\, a(y, \delta, l)\, \chi_1(y_1)\, \chi_1(y_2)\, dy. \tag{5-16}$$

One can consider either the substitution $(z_1, z_2) = (\lambda_1 y_1, \lambda_2 2^{-l} y_2)$, or the substitution $(z_1, z_2) = (\lambda_1 y_1, \lambda_3 \phi^a(y, \delta, l))$ (in order to carry this out one needs to consider the cases $y_2 \sim 1$ and $y_2 \sim -1$ separately), and get

$$\|\nu_j^\lambda\|_{L^\infty} \lesssim \min\{2^l \lambda_3, \lambda_2\}.$$

But now since $\lambda_2 \ll 2^l$ we may take $A_2 := \lambda_2$.

It remains to calculate the L^∞ bound for the $\hat{\nu}_l^\lambda$ function. This we can do by estimating the oscillatory integral in (5-12). As the calculations for the oscillatory integral in this case are almost identical to the ones in [Palle 2021, Section 5.5], we shall only briefly explain the case when $\lambda_1 \sim \lambda_2$, $2^{-l}\lambda_2 \ll \lambda_3 \ll \lambda_2$, corresponding to Case 6 in [Palle 2021, Section 5.5]. In all the other cases one gets that one can sum absolutely in the operator norm the operator pieces T_l^λ.

Let us remark that since $\lambda_2 \ll 2^l$, the case when $\lambda_1 \sim \lambda_2$, $2^{-l}\lambda_2 \sim \lambda_3$, corresponding to Case 4 in [Palle 2021, Section 5.5], does not appear anymore. This is critical since in this case one would not have absolute summability, nor would the complex interpolation method developed in [Ikromov and Müller 2016] work. This is the reason why we needed to consider T_I and T_{II} separately.

Case $\lambda_1 \sim \lambda_2$ and $2^{-l}\lambda_2 \ll \lambda_3 \ll \lambda_2$. As was obtained in [Palle 2021, Section 5.5], we have

$$\|\hat{\nu}_l^\lambda\|_{L^\infty} \lesssim \lambda_1^{-1/2} \lambda_3^{-N} \tag{5-17}$$

for any $N > 0$, that is, we have $A_1 = \lambda_1^{-1/2}\lambda_3^{-N}$, and recall that $A_2 = \lambda_2$, Therefore (5-15) gives

$$\|T_l^\lambda\|_{L^{\tilde{p}_3}_{x_3}(L^{\tilde{p}_1}_{(x_1,x_2)}) \to L^{\tilde{p}'_3}_{x_3}(L^{\tilde{p}'_1}_{(x_1,x_2)})} \lesssim \lambda_3^{-N}.$$

In order to be able to sum in $\lambda_1 \sim \lambda_2$ we need to use the complex interpolation method from [Ikromov and Müller 2016]. For a fixed λ_3 and ζ a complex number we define the measure $\mu_\zeta^{\lambda_3}$ by

$$\mu_\zeta^{\lambda_3} := \gamma(\zeta) \sum_{\lambda_1,\lambda_2} \lambda_1^{(1-3\zeta)/2} \nu_l^\lambda,$$

where the sum is over $\lambda_3 \ll \lambda_2 \ll 2^l$ and $\lambda_1 \sim \lambda_2$, and where $\gamma(\zeta) = 2^{-3(\zeta-1)/2} - 1$. We denote the associated convolution operator by $T_\zeta^{\lambda_3}$ and we recover with $\zeta = \frac{1}{3}$ the operator we want to estimate.

By a complex interpolation argument it suffices to show that

$$\begin{aligned}
\|T_{it}^{\lambda_3}\|_{L_{x_3}^{p_3}(L_{(x_1,x_2)}^{p_1}) \to L_{x_3}^{p_3'}(L_{(x_1,x_2)}^{p_1'})} &\lesssim \lambda_3^{-N} \quad \text{for } \left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(0, \frac{1}{4}\right),\\
\|T_{1+it}^{\lambda_3}\|_{L_{x_3}^{p_3}(L_{(x_1,x_2)}^{p_1}) \to L_{x_3}^{p_3'}(L_{(x_1,x_2)}^{p_1'})} &\lesssim 1 \quad\quad \text{for } \left(\frac{1}{p_1'}, \frac{1}{p_3'}\right) = \left(\frac{1}{2}, \frac{1}{2}\right)
\end{aligned}$$

for some $N > 0$, with constants uniform in $t \in \mathbb{R}$. The first estimate follows directly from the fact that $\hat{\nu}_l^\lambda$ have essentially disjoint supports with respect to λ and the estimate (5-17) (see [Palle 2021, Lemma 3.8(i)]), and for the other bound we need to estimate the L^∞ norm of the corresponding sum of the expressions (5-16). The proof is the same as in [Palle 2021, Section 5.5, Case 6], up to the formal difference in the function ϕ^a, which here behaves like y_2^k, and there like y_2^2. Since the domain of integration in (5-16) is $|y_2| \sim 1$, this is not essential. This finishes (the sketch of) the proof of the Fourier restriction for the nonadapted case, and also the proof of Theorem 1.2.

Appendix: Application of the Christ–Kiselev lemma

Recall that we consider the nonhomogeneous initial problem

$$\begin{cases} (\partial_t - i\phi(D))u(x,t) = F(x,t), & (x,t) \in \mathbb{R}^2 \times (0,\infty), \\ u(x,0) = G(x), & x \in \mathbb{R}^2, \end{cases}$$

for $F \in \mathcal{S}(\mathbb{R}^3)$, $G \in \mathcal{S}(\mathbb{R}^2)$, where ϕ, $\mathcal{W}$, and $(p_1, p_3) \in (1,2)^2$ are either as in Theorem 1.1 or 1.2, and where we additionally assume $\rho \in \{0, 1\}$. Note that ϕ is locally bounded and has polynomial growth at infinity, and note that according to Remark 2.3 the weight $\mathcal{W}$ is locally integrable in $\mathbb{R}^2$. The formula for a solution of the above equation is obtained through the Duhamel principle:

$$u(x,t) = (e^{i\phi(D)t} G)(x) + \int_0^t (e^{i\phi(D)(t-s)} F(\cdot, s))(x,s)\, \mathrm{d}s. \tag{A-1}$$

Note that $u \in C^\infty(\mathbb{R}^2 \times \mathbb{R}) \cap L_t^\infty((C_0)_{(x_1,x_2)}(\mathbb{R}^2))$, where C_0 denotes the space of continuous functions which tend to 0 at infinity.

We consider the following two surface measures (the second defined as in (1-3)):

$$\langle \mu_\phi, f\rangle = \int_{\mathbb{R}^2 \setminus \{0\}} f(x_1, x_2, \phi(x_1,x_2))\, \mathrm{d}x,$$

$$\langle \mu, f\rangle = \int_{\mathbb{R}^2 \setminus \{0\}} f(x_1, x_2, \phi(x_1,x_2))\, \mathcal{W}(x_1,x_2)\, \mathrm{d}x,$$

and we assume that the Fourier restriction estimate (1-2) for μ holds true for $(p_1, p_3) \in (1,2)^2$. One can easily check that

$$(e^{i\phi(D)t}G)(x) = \mathcal{F}^{-1}((\mathcal{F}G)\,\mathrm{d}\mu_\phi)(x,t) = \mathcal{F}^{-1}(\mathcal{W}^{-1}(\mathcal{F}G)\,\mathrm{d}\mu)(x,t),$$

and so this is precisely the Fourier extension operator of μ applied to the function $\mathcal{W}^{-1}\mathcal{F}G$. We can therefore bound the $L_t^{p_3'}(L_{(x_1,x_2)}^{p_1'})$ norm of this expression by the $L^2(\mathrm{d}\mu)$ norm of $\mathcal{W}^{-1}\mathcal{F}G$.

It remains to estimate the $L_t^{p_3'}(L_{(x_1,x_2)}^{p_1'})$ norm of the second term in (A-1). It turns out that the operator associated to this second term is closely related to the operator $f \mapsto f * \mathcal{F}^{-1}\mu$ (which we know is bounded from $L_t^{p_3}(L_{(x_1,x_2)}^{p_1})$ to $L_t^{p_3'}(L_{(x_1,x_2)}^{p_1'})$ since this is the corresponding R^*R operator). Namely, one can check that

$$\int_0^\infty (e^{i\phi(D)(t-s)}F(\cdot\,,s))(x,s)\,\mathrm{d}s = ((F\chi_{(0,\infty)}(s)) * (\mathcal{F}^{-1}\mu_\phi))(x,t),$$

and therefore it remains to pass from μ_ϕ to μ and to pass from integrating over $(0,\infty)$ in s to integrating over $(0,t)$ in s.

In order to do this, our first step is to use the Littlewood–Paley theorem in the x-direction so that our problem is reduced to proving the boundedness of the operator

$$\int_0^t (e^{i\phi(D)(t-s)}\,\eta_j(D)\,F(\cdot\,,s))(x,s)\,\mathrm{d}s, \tag{A-2}$$

where $(\eta_j)_{j\in\mathbb{Z}}$, $\eta_j = \eta \circ \delta_{2^{-j}}$, constitutes a partition of unity in $\mathbb{R}^2 \setminus \{0\}$ (as in (2-1) in Section 2A) respecting the α-mixed homogeneous dilation $\delta_{2^{-j}}$ defined in (1-1). By unwinding the definition of the operator in (A-2) and inserting the $\mathcal{W}$-factor, one obtains the expression (up to a universal constant)

$$\int_0^t \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} e^{i(x-y)\cdot\xi + i(t-s)\phi(\xi)} \eta_j(\xi)\mathcal{W}(\xi)\,\mathrm{d}\xi \right) F_{\mathcal{W}^{-1}}(y,s)\,\mathrm{d}y\,\mathrm{d}s, \tag{A-3}$$

where $F_{\mathcal{W}^{-1}} = \mathcal{F}^{-1}_{(x_1,x_2)}(\mathcal{W}^{-1}\mathcal{F}_{(x_1,x_2)}F)$. The expression within the brackets defines a convolution kernel $K_j(t-s; x-y)$ whose associated operator $T_j(t-s)$ in the x-variable is a bounded mapping from $L^{p_0}(\mathbb{R}^2)$ to $L^{p_0'}(\mathbb{R}^2)$ for any $p_0 \in [1,2]$ (since the integrand in the brackets is an $L_c^\infty(\mathbb{R}^2)$ function). Using the dominated convergence theorem one can get strong continuity of the operator-valued function $T_j : \mathbb{R} \to \mathcal{L}(L^{p_0}(\mathbb{R}^2); L^{p_0'}(\mathbb{R}^2))$ (which in turn, by the uniform boundedness principle, implies joint continuity $T_j : \mathbb{R} \times L^{p_0}(\mathbb{R}^2) \to L^{p_0'}(\mathbb{R}^2)$).

We may now apply the Christ–Kiselev lemma (for a proof of this variant see, e.g., [Sogge 1995, Chapter IV, Lemma 2.1]):

Lemma A.1. *Let Y and Z be separable Banach spaces and let $K : \mathbb{R} \to \mathcal{L}(Y, Z)$ be a continuous function from the real numbers to the space of bounded linear mappings $Y \to Z$ equipped with the strong operator topology. If the operator defined by*

$$(Tf)(t) := \int_{\mathbb{R}} K(t-s)\,f(s)\,\mathrm{d}s$$

is a bounded mapping from $L^{p_0}(\mathbb{R}, Y)$ *to* $L^{p_0'}(\mathbb{R}, Z)$ *for some* $p_0 \in (1, 2)$, *then the operator defined by*

$$(Wf)(t) := \int_{-\infty}^{t} K(t-s)\, f(s)\, \mathrm{d}s$$

is also a bounded mapping from $L^{p_0}(\mathbb{R}, Y)$ *to* $L^{p_0'}(\mathbb{R}, Z)$, *and in particular*

$$\|W\|_{L^{p_0}(\mathbb{R},Y)\to L^{p_0'}(\mathbb{R},Z)} \lesssim_{p_0} \|T\|_{L^{p_0}(\mathbb{R},Y)\to L^{p_0'}(\mathbb{R},Z)}.$$

Then we get that the $L_t^{p_3}(L^{p_1}_{(x_1,x_2)}) \to L_t^{p_3'}(L^{p_1'}_{(x_1,x_2)})$ boundedness of the operator in (A-3) (acting on $F_{\mathcal{W}^{-1}}$) is implied by the $L_t^{p_3}(L^{p_1}_{(x_1,x_2)}) \to L_t^{p_3'}(L^{p_1'}_{(x_1,x_2)})$ boundedness of the operator

$$\int_0^\infty \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}^2} e^{i(x-y)\cdot\xi + i(t-s)\phi(\xi)} \eta_j(\xi) \mathcal{W}(\xi)\, \mathrm{d}\xi \right) F_{\mathcal{W}^{-1}}(y,s)\, \mathrm{d}y\, \mathrm{d}s = ((F_{\mathcal{W}^{-1}} \chi_{(0,\infty)}(s)) * (\mathcal{F}^{-1}\mu_j))(x,t),$$

with essentially the same operator constant bound (up to a multiplicative factor which depends only on $p_3 \in (1,2)$). Here μ_j is the localized measure defined in the same way as in (2-2), and recall that this convolution operator is bounded (uniformly in j). This finishes the proof of Corollary 1.5.

References

[Abi-Khuzam and Shayya 2006] F. Abi-Khuzam and B. Shayya, "Fourier restriction to convex surfaces of revolution in $\mathbb{R}^3$", *Publ. Mat.* **50**:1 (2006), 71–85. MR Zbl

[Carbery and Ziesler 2002] A. Carbery and S. Ziesler, "Restriction and decay for flat hypersurfaces", *Publ. Mat.* **46**:2 (2002), 405–434. MR Zbl

[Carbery et al. 2007] A. Carbery, C. Kenig, and S. Ziesler, "Restriction for flat surfaces of revolution in $\mathbb{R}^3$", *Proc. Amer. Math. Soc.* **135**:6 (2007), 1905–1914. MR Zbl

[Carbery et al. 2013] A. Carbery, C. E. Kenig, and S. N. Ziesler, "Restriction for homogeneous polynomial surfaces in $\mathbb{R}^3$", *Trans. Amer. Math. Soc.* **365**:5 (2013), 2367–2407. MR Zbl

[Christ and Kiselev 2001] M. Christ and A. Kiselev, "Maximal functions associated to filtrations", *J. Funct. Anal.* **179**:2 (2001), 409–425. MR Zbl

[Cowling et al. 1990] M. Cowling, S. Disney, G. Mauceri, and D. Müller, "Damping oscillatory integrals", *Invent. Math.* **101**:2 (1990), 237–260. MR Zbl

[Dendrinos and Zimmermann 2019] S. Dendrinos and E. Zimmermann, "On L^p-improving for averages associated to mixed homogeneous polynomial hypersurfaces in $\mathbb{R}^3$", *J. Anal. Math.* **138**:2 (2019), 563–595. MR Zbl

[Drury 1990] S. W. Drury, "Degenerate curves and harmonic analysis", *Math. Proc. Cambridge Philos. Soc.* **108**:1 (1990), 89–96. MR Zbl

[Fefferman 1970] C. Fefferman, "Inequalities for strongly singular convolution operators", *Acta Math.* **124** (1970), 9–36. MR Zbl

[Ferreyra and Urciuolo 2009] E. Ferreyra and M. Urciuolo, "Fourier restriction estimates to mixed homogeneous surfaces", *J. Inequal. Pure Appl. Math.* **10**:2 (2009), art. id. 35. MR Zbl

[Ferreyra et al. 2004] E. Ferreyra, T. Godoy, and M. Urciuolo, "Restriction theorems for the Fourier transform to homogeneous polynomial surfaces in $\mathbb{R}^3$", *Studia Math.* **160**:3 (2004), 249–265. MR Zbl

[Ginibre and Velo 1992] J. Ginibre and G. Velo, "Smoothing properties and retarded estimates for some dispersive evolution equations", *Comm. Math. Phys.* **144**:1 (1992), 163–188. MR Zbl

[Greenblatt 2018] M. Greenblatt, "Fourier transforms of irregular mixed homogeneous hypersurface measures", *Math. Nachr.* **291**:7 (2018), 1075–1087. MR Zbl

[Gressman 2016] P. T. Gressman, "Damping oscillatory integrals by the Hessian determinant via Schrödinger", *Math. Res. Lett.* **23**:2 (2016), 405–430. MR Zbl

[Ikromov and Müller 2011] I. A. Ikromov and D. Müller, "On adapted coordinate systems", *Trans. Amer. Math. Soc.* **363**:6 (2011), 2821–2848. MR Zbl

[Ikromov and Müller 2016] I. A. Ikromov and D. Müller, *Fourier restriction for hypersurfaces in three dimensions and Newton polyhedra*, Ann. of Math. Stud. **194**, Princeton Univ. Press, 2016. MR Zbl

[Ikromov et al. 2005] I. A. Ikromov, M. Kempe, and D. Müller, "Damped oscillatory integrals and boundedness of maximal operators associated to mixed homogeneous hypersurfaces", *Duke Math. J.* **126**:3 (2005), 471–490. MR Zbl

[Iosevich and Sawyer 1996] A. Iosevich and E. Sawyer, "Oscillatory integrals and maximal averages over homogeneous surfaces", *Duke Math. J.* **82**:1 (1996), 103–141. MR Zbl

[Keel and Tao 1998] M. Keel and T. Tao, "Endpoint Strichartz estimates", *Amer. J. Math.* **120**:5 (1998), 955–980. MR Zbl

[Kenig et al. 1991] C. E. Kenig, G. Ponce, and L. Vega, "Oscillatory integrals and regularity of dispersive equations", *Indiana Univ. Math. J.* **40**:1 (1991), 33–69. MR Zbl

[Lizorkin 1970] P. I. Lizorkin, "Multipliers of Fourier integrals and bounds of convolution in spaces with mixed norms: applications", *Izv. Akad. Nauk SSSR Ser. Mat.* **34**:1 (1970), 218–247. In Russian; translated in *Math. USSR-Izv.* **4**:1 (1970), 225–255. MR Zbl

[Milnor 1964] J. Milnor, "On the Betti numbers of real varieties", *Proc. Amer. Math. Soc.* **15**:2 (1964), 275–280. MR Zbl

[Nomizu and Sasaki 1994] K. Nomizu and T. Sasaki, *Affine differential geometry*: *geometry of affine immersions*, Cambridge Tracts Math. **111**, Cambridge Univ. Press, 1994. MR Zbl

[Oberlin 2004] D. M. Oberlin, "A uniform Fourier restriction theorem for surfaces in $\mathbb{R}^3$", *Proc. Amer. Math. Soc.* **132**:4 (2004), 1195–1199. MR Zbl

[Oberlin 2012] D. M. Oberlin, "A uniform Fourier restriction theorem for surfaces in $\mathbb{R}^d$", *Proc. Amer. Math. Soc.* **140**:1 (2012), 263–265. MR Zbl

[Palle 2021] L. Palle, "Mixed norm Strichartz-type estimates for hypersurfaces in three dimensions", *Math. Z.* **297**:3-4 (2021), 1529–1599. MR Zbl

[Schwend 2020] J. Schwend, "Optimal $L^p \to L^q$ estimates for Euclidean averages over prototypical hypersurfaces in $\mathbb{R}^3$", preprint, 2020. arXiv 2012.15789

[Sjölin 1974] P. Sjölin, "Fourier multipliers and estimates of the Fourier transform of measures carried by smooth curves in R^2", *Studia Math.* **51** (1974), 169–182. MR Zbl

[Sogge 1995] C. D. Sogge, *Lectures on nonlinear wave equations*, Monogr. Anal. **2**, International, Boston, 1995. MR Zbl

[Strichartz 1977] R. S. Strichartz, "Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations", *Duke Math. J.* **44**:3 (1977), 705–714. MR Zbl

[Su 1983] Su B., *Affine differential geometry*, Sci. Press, Beijing, 1983. MR Zbl

[Triebel 2006] H. Triebel, *Theory of function spaces, III*, Monogr. Math. **100**, Birkhäuser, Basel, 2006. MR Zbl

[Varchenko 1976] A. N. Varchenko, "Newton polyhedra and estimation of oscillating integrals", *Funkt. Anal. i Prilozhen.* **10**:3 (1976), 13–38. In Russian; translated in *Funct. Anal. Appl.* **10**:3 (1976), 175–196. MR Zbl

[Zygmund 1974] A. Zygmund, "On Fourier coefficients and transforms of functions of two variables", *Studia Math.* **50** (1974), 189–201. MR Zbl

Received 19 Apr 2020. Revised 27 Mar 2021. Accepted 19 May 2021.

LJUDEVIT PALLE: palle@math.uni-kiel.de
Christian-Albrechts-Universität zu Kiel, Kiel, Germany

ANALYSIS AND PDE
Vol. 16 (2023), No. 1, pp. 233–277
DOI: 10.2140/apde.2023.16.233

HOLOMORPHIC FACTORIZATION OF MAPPINGS INTO $\mathbf{Sp_4}(\mathbb{C})$

BJÖRN IVARSSON, FRANK KUTZSCHEBAUCH AND ERIK LØW

We prove that any null-homotopic holomorphic map from a Stein space X to the symplectic group $\mathrm{Sp}_4(\mathbb{C})$ can be written as a finite product of elementary symplectic matrices with holomorphic entries.

1. Introduction

The continuous or holomorphic parameter dependence of classical linear algebra results over the fields $\mathbb{R}$ or $\mathbb{C}$ form a circle of very natural questions of general mathematical interest. For example the factorization of continuous matrices as a product of continuous elementary matrices has been studied and solved by Vaserstein [1988]. The corresponding holomorphic problem for the special linear group SL_n was posed by Gromov [1989] and was finally solved by the first two authors in [Ivarsson and Kutzschebauch 2012]. The study of algebraic dependence is connected with famous work by Suslin [1977], Cohn [1966], Bass, Milnor, and Serre [Bass et al. 1967] and many others.

Part of this research was done while the authors were visitors at The Centre for Advanced Study (CAS) at the Norwegian Academy of Science and Letters. Ivarsson was also supported by the Magnus Ehrnrooth Foundation and Løw by Bergens Forskningsstiftelse (BFS). The research of Kutzschebauch was partially supported by Schweizerische Nationalfonds Grant 200021-178730.
MSC2020: primary 32Q56; secondary 19B14.
Keywords: Oka principle, Stein spaces, symplectic matrices, unipotent factorization, elementary symplectic matrices, rings of holomorphic functions, algebraic K-theory.

These parameter dependence questions are a part of algebraic K-theory and the study of linear algebra over general rings. Factorization of Chevalley groups over $\mathbb{R}$ and $\mathbb{C}$ into elementary matrices is classically well known. For Chevalley groups over general rings this is much more difficult and studied a lot. For an overview, see, for example, [Vavilov and Stepanov 2011].

We are mainly interested in the rings of holomorphic functions on Stein spaces. The only known holomorphic result is the existence for the special linear groups in [Ivarsson and Kutzschebauch 2012], where Gromov's problem is solved in full generality. In the special case of an open Riemann surface the problem was solved earlier (absolutely unnoticed) by Klein and Ramspott [1987]. The authors [Ivarsson et al. 2020] also proved the main result of this paper for any size of symplectic matrices in the special case of an open Riemann surface.

In the present paper we consider the symplectic groups over rings of holomorphic functions on Stein spaces. The main result is (see Section 2 for notation)

Main Theorem (Theorem 3.1). *Let X be a finite-dimensional reduced Stein space and $f : X \to \mathrm{Sp}_4(\mathbb{C})$ be a holomorphic mapping that is null-homotopic. Then there exist a natural number K, depending only on the dimension of X, and holomorphic mappings*

$$G_1, \dots, G_K : X \to \mathbb{C}^3$$

such that

$$f(x) = M_1(G_1(x)) \cdots M_K(G_K(x)).$$

We remind the reader that a mapping is null-homotopic if it is homotopic to a constant map. By the Oka–Grauert principle it is equivalent for a holomorphic map from a Stein space into a complex Lie group to be null-homotopic via holomorphic maps or via continuous maps (see Theorem 5.3.2 in the standard reference [Forstnerič 2017]).

Our main tool is the Oka principle for stratified elliptic submersions, the most elaborate result in modern Oka theory. In order to apply an Oka principle one needs a topological solution which we take from our previous work on symplectic groups over rings of continuous functions on topological spaces. The Oka principle lets us homotope the topological solution to a holomorphic one. The technical details needed to prove that certain fibrations are stratified elliptic are considerable and we have so far only been able to complete these details for Sp_4. We expect that a similar result holds for Sp_{2n}.

Factorization of symplectic groups over other rings (of mainly algebraic nature) has been considered before for example by Kopeiko [1978], and Grunewald, Mennicke and Vaserstein [Grunewald et al. 1991].

The paper is organized as follows. In Section 2 we recall our results on factorization of continuous matrices and prove a slight extension about the number of factors. In Section 3 we state our main results and give an overview over the proof. In Section 4 we explain how our results can be reformulated in the language used in algebraic K-theory. In Section 5 we recall the theorems from Oka theory which we use in our proof.

In Section 6 we give the proofs of Lemmas 3.3 and 3.4, where we prove that the most important fibrations in this paper, the projections of products of elementary symplectic matrices onto their last row, are surjective and we determine where they are submersive. This is done for symplectic matrices of all sizes, since we hope to be able to prove in the future that these fibrations are stratified elliptic for all sizes.

The rest of the paper is devoted to proving that our fibration (for (4×4)-matrices) is stratified elliptic in order to be able to apply Oka theory. In Section 7 we describe the stratification with respect to which we want to prove that the important fibration is stratified elliptic. This has to do with how the set of $2n$ algebraic equations defining a fiber in the fibration can be reduced to n equations. In the case of the special linear group in [Ivarsson and Kutzschebauch 2012] we were able to reduce to one single equation independent of the size of the matrices. This was the crucial trick to prove ellipticity by finding complete vector fields, which corresponds to Gromov's example of a spray. This inability to reduce to fewer equations is the main difference between the situation of the symplectic group and the special linear group. It leads to all the difficult technical work in the rest of the paper. In Section 8 we introduce our method to find complete vector fields tangent to the fibration. However not all of them are complete and we deduce that the Gromov-spray produced by them is not dominating. We determine which of them are complete. In Section 9 we explain our strategy to enlarge the set of complete vector fields so that this enlarged collection now spans the tangent space at all points and thus gives a fiber-dominating spray. The realization of this strategy takes Section 10, where we introduce useful quantities, Sections 11, 12, and 13, where we prove the result for three, four, and five (elementary symplectic) factors, and finally we can give an inductive (over the number of factors) proof in Section 14. The reason for dealing with the low numbers of factors separately is that some of the fibers of our fibration are reducible when there are a small number of factors, and from five factors on all fibers are irreducible. In Section 15 we end the paper with an application to the problem of a product of exponentials and formulate some open questions.

2. Continuous factorization

Let

$$\omega=\sum_{j=1}^{n} dz_j\wedge dz_{j+n}$$

be the symplectic form in $\mathbb{C}^{2n}$. With respect to ω, symplectic matrices are those that can be written in block form as

$$\begin{pmatrix} A & B\\ C & D\end{pmatrix},$$

where A, B, C and D are complex $n\times n$ matrices satisfying

$$A^TC=C^TA, \tag{2.0.1}$$

$$B^TD=D^TB, \tag{2.0.2}$$

$$A^TD-C^TB=I_n, \tag{2.0.3}$$

where I_n is the $n\times n$ identity matrix. In the case $B=C=0$ this means $D=(A^T)^{-1}$, and in the case $A=D=I_n$ this means B and C are symmetric and $C^TB=0$. Let U_n denote an $n\times n$ matrix satisfying $U_n=U_n^T$ and 0_n the $n\times n$ zero matrix. We call those matrices that are written in block form as

$$\begin{pmatrix} I_n & 0_n\\ U_n & I_n\end{pmatrix}\quad\text{or}\quad\begin{pmatrix} I_n & U_n\\ 0_n & I_n\end{pmatrix}$$

elementary symplectic matrices. Let

$$U_n(x_1,\dots,x_{n(n+1)/2}) = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ x_2 & x_{n+1} & \dots & x_{2n-1} \\ \vdots & \vdots & \ddots & \vdots \\ x_n & x_{2n-1} & \dots & x_{n(n+1)/2} \end{pmatrix}.$$

Given a map $G : X \to \mathbb{C}^{n(n+1)/2}$ let

$$U_n(G(x)) = U_n(G_1(x),\dots,G_{n(n+1)/2}(x)),$$

where the G_j's are components of the map G. For odd k let

$$M_k(G(x)) = \begin{pmatrix} I_n & 0_n \\ U_n(G(x)) & I_n \end{pmatrix},$$

and for even k

$$M_k(G(x)) = \begin{pmatrix} I_n & U_n(G(x)) \\ 0_n & I_n \end{pmatrix}.$$

The following result is a refinement of [Ivarsson et al. 2020, Theorem 1.3].

Theorem 2.1 (continuous Vaserstein problem for symplectic matrices). *There exists a natural number $K(n, d)$ such that given any finite-dimensional normal topological space X of (covering) dimension d and any null-homotopic continuous mapping $M : X \to \mathrm{Sp}_{2n}(\mathbb{C})$ there exist K continuous mappings*

$$G_1,\dots,G_K : X \to \mathbb{C}^{n(n+1)/2}$$

such that

$$M(x) = M_1(G_1(x)) \cdots M_K(G_K(x)).$$

Proof. Theorem 1.3 in [Ivarsson et al. 2020] does not give a uniform bound on the number of factors depending on n and d. Suppose such a bound does not exist; i.e., for all natural numbers i there are normal topological spaces X_i of dimension d and null-homotopic continuous maps $f_i : X_i \to \mathrm{Sp}_{2n}(\mathbb{C})$ such that f_i does not factor over a product of less than i elementary symplectic matrices. Let X equal $\bigcup_{i=1}^{\infty} X_i$, the disjoint union of the spaces X_i, and $F : X \to \mathrm{Sp}_{2n}(\mathbb{C})$ be the map that is equal to f_i on X_i. By Theorem 1.3 in [Ivarsson et al. 2020] F factors over a finite number of elementary symplectic matrices. Consequently all f_i factor over the same number of elementary symplectic matrices, which contradicts the assumption on f_i. □

3. Statement of the main result and overview of proof

We state the main result of this paper which is a holomorphic version of Theorem 2.1 for $\mathrm{Sp}_4(\mathbb{C})$.

Theorem 3.1. *There exists a natural number $N(d)$ such that given any finite-dimensional reduced Stein space X of dimension d and any null-homotopic holomorphic mapping $f : X \to \mathrm{Sp}_4(\mathbb{C})$ there exist N holomorphic mappings*

$$G_1,\dots,G_N : X \to \mathbb{C}^3$$

such that

$$f(x) = M_1(G_1(x)) \cdots M_N(G_N(x)).$$

We have the following corollary.

Corollary 3.2. *Let X be a finite-dimensional reduced Stein space that is topologically contractible and $f : X \to \mathrm{Sp}_4(\mathbb{C})$ be a holomorphic mapping. Then there exist a natural number N and holomorphic mappings*

$$G_1, \dots, G_N : X \to \mathbb{C}^3$$

such that

$$f(x) = M_1(G_1(x)) \cdots M_N(G_N(x)).$$

The strategy for proving Theorem 3.1 is as follows. Define

$$\Psi_K : (\mathbb{C}^3)^K \to \mathrm{Sp}_4(\mathbb{C})$$

as

$$\Psi_K(x_1, \dots, x_{3K}) = M_1(x_1, x_2, x_3) \cdots M_K(x_{3K-2}, x_{3K-1}, x_{3K}). \tag{3.0.1}$$

We want to show the existence of a holomorphic map

$$G = (G_1, \dots, G_K) : X \to (\mathbb{C}^3)^K$$

such that

$$\begin{tikzcd} & (\mathbb{C}^3)^K \arrow[d, "\Psi_K"] \\ X \arrow[ur, "G"] \arrow[r, "f"'] & \mathrm{Sp}_4(\mathbb{C}) \end{tikzcd}$$

is commutative. Theorem 2.1 shows the existence of a continuous map such that the diagram above is commutative.

We will prove Theorem 3.1 using the Oka–Grauert–Gromov principle for sections of holomorphic submersions over X. One candidate submersion would be to use the pull-back of $\Psi_K : (\mathbb{C}^3)^K \to \mathrm{Sp}_4(\mathbb{C})$. It turns out that Ψ_K is not a submersion at all points in $(\mathbb{C}^3)^K$. It is a surjective holomorphic submersion if one removes a certain subset from $(\mathbb{C}^3)^K$. Unfortunately the fibers of this submersion are quite difficult to analyze and we therefore elect to study

$$\begin{tikzcd} & (\mathbb{C}^3)^K \arrow[d, "\pi_4 \circ \Psi_K"] \\ X \arrow[ur, "F"] \arrow[r, "\pi_4 \circ f"'] & \mathbb{C}^4 \setminus \{0\} \end{tikzcd}$$

where we define the projection $\pi_4 : \mathrm{Sp}_4(\mathbb{C}) \to \mathbb{C}^4 \setminus \{0\}$ to be the projection of a matrix to its last row:

$$\pi_4 \begin{pmatrix} z_{11} & \dots & z_{14} \\ \vdots & \ddots & \vdots \\ z_{41} & \dots & z_{44} \end{pmatrix} = (z_{41}, \dots, z_{44}).$$

However, even the map $\Phi_K = \pi_4 \circ \Psi_K : (\mathbb{C}^3)^K \to \mathbb{C}^4 \setminus \{0\}$ is not submersive everywhere. We have the three results below (Lemmas 3.3 and 3.4 and Proposition 3.6) about that map which will be proved in later sections.

We introduce some notation. Projecting to the last row introduces an asymmetry between upper and lower triangular elementary matrices and therefore we will denote by z's the variables in the lower triangular matrices and by w's the variables in the upper triangular matrices. For example, the right-hand side of (3.0.1) becomes

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ z_1 & z_2 & 1 & 0 \\ z_2 & z_3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & w_1 & w_2 \\ 0 & 1 & w_2 & w_3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & 0 & w_{3k-2} & w_{3k-1} \\ 0 & 1 & w_{3k-1} & w_{3k} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

for even $K = 2k$.

Let

$$\vec{Z}_K = \begin{cases} (z_1, z_2, z_3, w_1, w_2, w_3, \ldots, w_{3k-2}, w_{3k-1}, w_{3k}) & \text{if } K = 2k, \\ (z_1, z_2, z_3, w_1, w_2, w_3, \ldots, z_{3k+1}, z_{3k+2}, z_{3k+3}) & \text{if } K = 2k+1 \end{cases}$$

and

$$W_K = \begin{cases} \begin{pmatrix} w_1 & w_2 & w_4 & w_5 & \ldots & w_{3k-5} & w_{3k-4} \\ w_2 & w_3 & w_5 & w_6 & \ldots & w_{3k-4} & w_{3k-3} \end{pmatrix} & \text{if } K = 2k, \\ \begin{pmatrix} w_1 & w_2 & w_4 & w_5 & \ldots & w_{3k-2} & w_{3k-1} \\ w_2 & w_3 & w_5 & w_6 & \ldots & w_{3k-1} & w_{3k} \end{pmatrix} & \text{if } K = 2k+1. \end{cases}$$

Also, when $K = 2k$ or $K = 2k+1$, let

$$A_K = \bigcap_{1 \le j \le k} \{\vec{Z}_K \in (\mathbb{C}^3)^K : z_{3j-1} = z_{3j} = 0\},$$

$$B_K = \{\vec{Z}_K \in (\mathbb{C}^3)^K : \text{Rank } W_K < 2\}$$

and

$$S_K = A_K \cap B_K. \tag{3.0.2}$$

We have Lemma 3.3, which follows from a simple calculation.

Lemma 3.3. *The mapping*

$$\Phi_K = \pi_4 \circ \Psi_K : (\mathbb{C}^3)^K \setminus S_K \to \mathbb{C}^4 \setminus \{0\}$$

is surjective when $K \ge 3$.

Lemma 3.4. *For $K \ge 3$ the mapping*

$$\Phi_K = \pi_4 \circ \Psi_K : (\mathbb{C}^3)^K \to \mathbb{C}^4 \setminus \{0\}$$

is a holomorphic submersion exactly at points $\vec{Z}_K \in (\mathbb{C}^3)^K \setminus S_K$, where S_K is defined by (3.0.2) above. That is, S_K is the set of points where the entries in the last row of each lower triangular matrix are zero, except for the K-th matrix where no conditions are imposed, and the rank of the matrix W_K, which does not involve entries from the K-th matrix, is strictly less than 2.

Remark 3.5. Lemmas 3.3 and 3.4 both generalize to $2n\times 2n$ matrices and the proofs are identical. In Section 6 we therefore consider the general case.

Proposition 3.6. *For $n=1$ and $n=2$ the map*

$$\begin{tikzcd}(\mathbb{C}^{n(n+1)/2})^K \setminus S_K \arrow[d, "\pi_{2n}\circ\Psi_K"] \\ \mathbb{C}^{2n} \setminus \{0\}\end{tikzcd} \tag{3.0.3}$$

is a stratified elliptic submersion.

Corollary 3.7. *Let $n=1$ or $n=2$. Let X be a finite-dimensional reduced Stein space and $f : X \to \mathrm{Sp}_{2n}(\mathbb{C})$ be a holomorphic map. Assume that there exists a natural number K and a continuous map $F : X \to (\mathbb{C}^{n(n+1)/2})^K \setminus S_K$ such that*

$$\begin{tikzcd} & (\mathbb{C}^{n(n+1)/2})^K \setminus S_K \arrow[d, "\pi_{2n}\circ\Psi_K"] \\ X \arrow[ur, "F"] \arrow[r, "\pi_{2n}\circ f"'] & \mathbb{C}^{2n} \setminus \{0\}\end{tikzcd}$$

is commutative. Then there exists a holomorphic map $G : X \to (\mathbb{C}^{n(n+1)/2})^K \setminus S_K$, homotopic to F via continuous maps $F_t : X \to (\mathbb{C}^{n(n+1)/2})^K \setminus S_K$, such that the diagram above is commutative for all F_t.

Proof. The pull-back of (3.0.3) by $\pi_{2n} \circ f$ is a stratified elliptic submersion over the Stein base X. Thus by Theorem 5.6 there is a homotopy from the given continuous section to a holomorphic section. This is equivalent to the desired homotopy F_t. An even better way to perform this proof is to say that the map (3.0.3) is an Oka map, see [Forstnerič 2017, Corollary 7.4.5(i)], which yields the desired conclusion. ☐

Remark 3.8. The fact that the map (3.0.3) is an Oka map yields a parametric version of Corollary 3.7. This means that the holomorphic map can be replaced by a continuous map $f_P : X \times P \to \mathrm{Sp}_{2n}(\mathbb{C})$, which is holomorphic for each fixed parameter $p \in P$, where P is a compact Hausdorff topological space.

We need the following version of the Whitehead lemma:

$$\begin{pmatrix}1&0&0&0\\a&1&0&0\\0&0&1&-a\\0&0&0&1\end{pmatrix} = \begin{pmatrix}1&0&0&0\\0&1&0&0\\-a&-1&1&0\\-1&0&0&1\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1&0&-a\\0&0&1&0\\0&0&0&1\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&1&1&0\\1&0&0&1\end{pmatrix}\begin{pmatrix}1&0&0&0\\0&1&0&a\\0&0&1&0\\0&0&0&1\end{pmatrix}. \tag{3.0.4}$$

Proof of Theorem 3.1. We will prove the theorem for a single map. The existence of a uniform bound $N(d)$ follows as in the proof of Theorem 2.1. Since a finite-dimensional Stein space is finite-dimensional as a topological space there are $K-2$ continuous mappings

$$G_1, \dots, G_{K-2} : X \to \mathbb{C}^3$$

such that

$$f(x) = M_1(G_1(x)) \cdots M_{K-2}(G_{K-2}(x)).$$

Let $H : X \to \mathbb{C}^3$ be a constant map such that $U_2(H)$ has nonzero second row, let $\mathbb{O} : X \to \mathbb{C}^3$ be the zero map, and replace the above factorization by

$$f(x) = M_1(H)M_2(\mathbb{O})M_3(G_1(x) - H)M_4(G_2(x)) \cdots M_K(G_{K-2}(x))$$

(suppressing the variables in the constant maps H and $\mathbb{O}$). Notice that the second factor is the identity matrix.

This factorization by K continuous elementary symplectic matrices avoids the singularity set S_K and thus we find $F : X \to (\mathbb{C}^3)^K \setminus S_K$ with $\Psi_K(F) = f$.

Using Corollary 3.7 we know that $F_0 := F$ is homotopic to a holomorphic map $G = F_1$, via continuous maps F_t, such that

$$\pi_4(f(x)) = \pi_4 \circ \Psi_K(F_t(x)), \quad 0 \le t \le 1,$$

that is, the last row of the matrices $\Psi_K(F_t(x))$ is constant. Therefore

$$\Psi_K(F_t(x))f(x)^{-1} = \begin{pmatrix} \tilde{f}_{11,t}(x) & \tilde{f}_{12,t}(x) & \tilde{f}_{13,t}(x) & \tilde{f}_{14,t}(x) \\ \tilde{f}_{21,t}(x) & \tilde{f}_{22,t}(x) & \tilde{f}_{23,t}(x) & \tilde{f}_{24,t}(x) \\ \tilde{f}_{31,t}(x) & \tilde{f}_{32,t}(x) & \tilde{f}_{33,t}(x) & \tilde{f}_{34,t}(x) \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Since these matrices are symplectic, it automatically follows that $\tilde{f}_{12,t}(x) \equiv 0$, $\tilde{f}_{22,t}(x) \equiv 1$, and $\tilde{f}_{32,t}(x) \equiv 0$ so that

$$\Psi_K(F_t(x))f(x)^{-1} = \begin{pmatrix} \tilde{f}_{11,t}(x) & 0 & \tilde{f}_{13,t}(x) & \tilde{f}_{14,t}(x) \\ \tilde{f}_{21,t}(x) & 1 & \tilde{f}_{23,t}(x) & \tilde{f}_{24,t}(x) \\ \tilde{f}_{31,t}(x) & 0 & \tilde{f}_{33,t}(x) & \tilde{f}_{34,t}(x) \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.0.5}$$

and in addition

$$\tilde{f}_t(x) = \begin{pmatrix} \tilde{f}_{11,t}(x) & \tilde{f}_{13,t}(x) \\ \tilde{f}_{31,t}(x) & \tilde{f}_{33,t}(x) \end{pmatrix} \in \mathrm{Sp}_2(\mathbb{C}) = \mathrm{SL}_2(\mathbb{C}). \tag{3.0.6}$$

Since $\Psi_K(F_0(x)) = f(x)$, we see that $\tilde{f}_0 = \mathrm{Id}$, and thus the holomorphic map $\tilde{f} := \tilde{f}_1 : X \to \mathrm{SL}_2(\mathbb{C})$ is null-homotopic. Let ψ be the standard inclusion of Sp_2 in Sp_4; see for example [Grunewald et al. 1991]. By the main result from [Ivarsson and Kutzschebauch 2012] the matrix

$$\psi(\tilde{f}(x)^{-1}) = \begin{pmatrix} \tilde{f}_{33}(x) & 0 & -\tilde{f}_{13}(x) & 0 \\ 0 & 1 & 0 & 0 \\ -\tilde{f}_{31}(x) & 0 & \tilde{f}_{11}(x) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.0.7}$$

is a product of holomorphic elementary symplectic matrices. Therefore it suffices to show that

$$\Psi_K(G(x))f(x)^{-1} \cdot \psi(\tilde{f}(x)^{-1}) = \begin{pmatrix} 1 & 0 & 0 & \tilde{f}_{14}(x) \\ -\tilde{f}_{34}(x) & 1 & \tilde{f}_{14}(x) & \tilde{f}_{24}(x) \\ 0 & 0 & 1 & \tilde{f}_{34}(x) \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.0.8}$$

is a product of elementary symplectic matrices. In order to deduce that the right-hand side of (3.0.8) has the claimed form one has to use the fact that (3.0.5) is symplectic. Since

$$\begin{pmatrix} 1 & 0 & 0 & \tilde{f}_{14}(x) \\ -\tilde{f}_{34}(x) & 1 & \tilde{f}_{14}(x) & \tilde{f}_{24}(x) \\ 0 & 0 & 1 & \tilde{f}_{34}(x) \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\tilde{f}_{34}(x) & 1 & 0 & 0 \\ 0 & 0 & 1 & \tilde{f}_{34}(x) \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & \tilde{f}_{14}(x) \\ 0 & 1 & \tilde{f}_{14}(x) & \tilde{f}_{14}(x)\tilde{f}_{34}(x)+\tilde{f}_{24}(x) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

the result follows by the Whitehead lemma, (3.0.4). □

Analyzing this proof and using Remark 3.8 one sees that we can actually prove a parametric version of our main theorem.

Theorem 3.9. *Let X be a finite-dimensional reduced Stein space, P be a compact Hausdorff topological (parameter) space, and $f : P \times X \to \mathrm{Sp}_4(\mathbb{C})$ be a continuous mapping, holomorphic for each fixed $p \in P$, that is null-homotopic. Then there exist a natural number K and continuous mappings, holomorphic for each fixed parameter $p \in P$,*

$$G_1, \dots, G_K : P \times X \to \mathbb{C}^3$$

such that

$$f(p, x) = M_1(G_1(p, x)) \cdots M_K(G_K(p, x)).$$

To complete the proof of the theorem we need to establish Proposition 3.6 and Lemmas 3.4 and 3.3.

Remark 3.10. Proposition 3.6 is the crucial ingredient in the proof of Theorem 3.1. Its proof is by far the most difficult part of the paper. As pointed out in Remark 3.5, Lemma 3.4 holds for general n. Also if Proposition 3.6 holds for some n then Corollary 3.7 also holds for that n. Moreover the reduction of the size of the symplectic matrix from Sp_4 to Sp_2 done in the proof of Theorem 3.1 generalizes easily to a reduction from Sp_{2n} to Sp_{2n-2} if Corollary 3.7 holds for n (see for example the proof of Lemma 4.4 in [Grunewald et al. 1991]). Therefore if Proposition 3.6 can be proven for $n = 1, \dots, m$ then the following holds true.

Conjecture 3.11. *Let X be a finite-dimensional reduced Stein space and $f : X \to \mathrm{Sp}_{2m}(\mathbb{C})$ be a holomorphic mapping that is null-homotopic. Then there exist a natural number K and holomorphic mappings*

$$G_1, \dots, G_K : X \to \mathbb{C}^{m(m+1)/2}$$

such that

$$f(x) = M_1(G_1(x)) \cdots M_K(G_K(x)).$$

In the case of a 1-dimensional Stein space, i.e., an open Riemann surface, this conjecture was established in [Ivarsson et al. 2020]. The condition of null-homotopy is automatically satisfied in this case, since an open Riemann surface is homotopy equivalent to a 1-dimensional CW-complex and the group $\mathrm{Sp}_{2m}(\mathbb{C})$ is simply connected. The proof uses the analytic ingredient that the Bass stable rank of $\mathcal{O}(X)$ is 1 for an open Riemann surface and proceeds then by linear algebra arguments.

4. Formulation in algebraic terms

We relate our results to algebraic K-theory and reformulate them in those terms. The following is a standard notion:

Definition 4.1. For a commutative ring R the set $U_m(R)$ of unimodular rows of length m is defined as

$$\{(r_1, r_2, \dots, r_m) \in R^m : r_1, r_2, \dots r_m \text{ generate } R \text{ as an ideal}\}.$$

In our main example, if $\mathcal{O}(X)$ is the ring of holomorphic functions on a Stein space X, a row $(f_1, f_2, \dots, f_m) \in \mathcal{O}^m(X)$ is unimodular if and only if the functions $f_1, f_2, \dots, f_m$ have no common zeros, a well-known application of Cartan's Theorem B.

Since null-homotopy is an important assumption in our studies we denote the set of null-homotopic unimodular rows in $U_m(\mathcal{O}(X))$ by $U_m^0(\mathcal{O}(X))$. This set can be seen as the path-connected component of the space of holomorphic maps from X to $\mathbb{C}^m \setminus \{0\}$ containing the constant map $(0, 0, , \dots, 0, 1) = e_m$. By the Oka–Grauert principle $\mathbb{C}^m \setminus \{0\} = \mathrm{GL}_m(\mathbb{C})/\mathrm{GL}_{m-1}(\mathbb{C})$ is an Oka manifold; therefore the path-connected components of continuous and holomorphic maps $X \to \mathbb{C}^m \setminus \{0\}$ are in bijection. This says that unimodular rows in $U_m(\mathcal{O}(X))$ are null-homotopic in the holomorphic sense if and only if they are null-homotopic in the continuous sense.

Algebraic K-theorists consider Chevalley groups over rings; in our example we consider the null-homotopic elements of them.

Definition 4.2. $\mathrm{Sp}_{2n}^0(\mathcal{O}(X))$ denotes the group of null-homotopic holomorphic maps from a Stein space X to the symplectic group $\mathrm{Sp}_{2n}(\mathbb{C})$, which in other words is the path-connected component of the group $\mathrm{Sp}_{2n}(\mathcal{O}(X))$ containing the identity.

Again by the Oka–Grauert principle holomorphic maps $X \to \mathrm{Sp}_{2n}(\mathbb{C})$ are homotopic via holomorphic maps if and only if they are homotopic via continuous maps.

Clearly the last row of a matrix in $\mathrm{Sp}_{2n}(\mathcal{O}(X))$ is unimodular, i.e., an element of $U_{2n}(\mathcal{O}(X))$. Whether a unimodular row in $U_{2n}(\mathcal{O}(X))$ is the last row of a matrix in $\mathrm{Sp}_{2n}(\mathcal{O}(X))$ is by Oka theory a purely topological problem. Let us illustrate this by an example.

Extending a unimodular row to an invertible matrix can be reformulated as follows: given a trivial line subbundle of the trivial bundle $X \times \mathbb{C}^n$ of rank n over X, can it be complemented by a trivial bundle?

This of course is not always the case: The (nontrivial) tangent bundle T of the sphere S^{2n+1} $(n \geq 4)$ is the complement of the trivial normal bundle N to the sphere S^{2n+1} in $\mathbb{R}^{2n+2}$. To make this a holomorphic example consider X to be a Grauert tube around S^{2n+1}, i.e., a Stein manifold which has a strong deformation retraction ρ onto its totally real maximal-dimensional submanifold S^{2n+1}. The bundle T is replaced by the complexified tangent bundle to the sphere pulled back onto X by the retraction ρ and equipped with its unique structure of holomorphic vector bundle (which is still not a trivial bundle). The pull-back of the complexified trivial bundle N is still a trivial line subbundle of $X \times \mathbb{C}^{2n}$. Thus we have found an example of a holomorphic row which cannot be completed to an invertible matrix in $\mathrm{GL}_{2n}(\mathcal{O}(X))$ and thus not to a matrix in $\mathrm{Sp}_{2n}(\mathcal{O}(X))$ either.

For null-homotopic rows the situation is better.

Lemma 4.3. *Every element $U^0_{2n}(\mathcal{O}(X))$ extends to a null-homotopic matrix $A \in \mathrm{Sp}^0_{2n}(\mathcal{O}(X))$.*

Proof. Let $F = (f_1, \dots, f_{2n}) : X \to \mathbb{C}^{2n} \setminus \{0\}$ be a null-homotopic holomorphic map, and the homotopy to the constant map $F_1(x) = e_{2n}$ be denoted by F_t, $t \in [0, 1]$. The map $\pi_{2n} : \mathrm{Sp}_{2n}(\mathbb{C}) \to \mathbb{C}^{2n} \setminus \{0\}$ is a locally trivial holomorphic fiber bundle with typical fiber $F \cong \mathrm{Sp}_{2n-2}(\mathbb{C}) \times \mathbb{C}^{4n-1}$ which is an Oka manifold. Our problem is to find a global section of the pull-back of this fibration by the map $F = F_0$. Since a locally trivial bundle is a Serre fibration and the constant last row can be extended to a constant (thus null-homotopic) symplectic matrix, we find a continuous section of this pull-back bundle over the whole homotopy. Thus the restriction to $X \times \{0\}$ is a null-homotopic continuous symplectic matrix. Since the fiber F is Oka, we find a homotopy to a holomorphic symplectic matrix, which is still null-homotopic. □

The notion of elementary symplectic matrices over a ring R is the same as explained in Section 2.

Let W_n denote an $n \times n$ matrix with entries in the ring R satisfying $W_n = W_n^T$ and 0_n the $n \times n$ zero matrix. We call those matrices that are written in block form as

$$\begin{pmatrix} I_n & 0_n \\ W_n & I_n \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} I_n & W_n \\ 0_n & I_n \end{pmatrix}$$

elementary symplectic matrices over R. The group generated by them, the elementary symplectic group, is denoted by $\mathrm{Ep}_{2n}(R)$. We consider the group $\mathrm{Ep}_{2n}(\mathcal{O}(X))$ which is easily seen to be a subgroup of $\mathrm{Sp}^0_{2n}(\mathcal{O}(X))$ (multiply the symmetric matrices W_n by a real number $t \in [0, 1]$).

The meaning of Corollary 3.7 in K-theoretic terms is now the following:

Proposition 4.4. *Let $n = 1$ or $n = 2$. For a Stein space X the group $\mathrm{Ep}_{2n}(\mathcal{O}(X))$ acts transitively on the set of null-homotopic unimodular rows $U^0_{2n}(\mathcal{O}(X))$.*

Proof. Let $u \in U^0_{2n}(\mathcal{O}(X))$ be a null-homotopic unimodular row. By the above lemma we can extend it to a null-homotopic symplectic matrix $A \in \mathrm{Sp}^0_{2n}(\mathcal{O}(X))$. Now we just follow the beginning of the proof of Theorem 3.1. By Theorem 2.1 we can factorize $A(x)$ as a product of elementary symplectic matrices with continuous entries. Adding two more elementary symplectic matrices we can achieve that the factorization avoids the singularity set S_K. Applying Corollary 3.7 we know that $A_0 := A$ is homotopic to a holomorphic map $G = A_1$, via continuous maps A_t, such that

$$\pi_4(A(x)) = \pi_4 \circ \Psi_K(A_t(x)), \quad 0 \le t \le 1,$$

that is, the last row of the matrices $\Psi_K(A_t(x))$ is constant. Therefore

$$\Psi_K(A_t(x))A(x)^{-1} = \begin{pmatrix} \tilde{a}_{11,t}(x) & \tilde{a}_{12,t}(x) & \tilde{a}_{13,t}(x) & \tilde{a}_{14,t}(x) \\ \tilde{a}_{21,t}(x) & \tilde{a}_{22,t}(x) & \tilde{a}_{23,t}(x) & \tilde{a}_{24,t}(x) \\ \tilde{a}_{31,t}(x) & \tilde{a}_{32,t}(x) & \tilde{a}_{33,t}(x) & \tilde{a}_{34,t}(x) \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

This shows that the element $\Psi_K(G(x))$ of $\mathrm{Ep}_{2n}(\mathcal{O}(X))$ has the last row equal to u or equivalently moves the constant row e_{2n} to u. □

Let $\psi : \mathrm{SL}_2 \to \mathrm{Sp}_4$ be the standard embedding given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a & 0 & b & 0 \\ 0 & 1 & 0 & 0 \\ c & 0 & d & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.0.1}$$

Continuing like in the proof of Theorem 3.1 we see that it gives the following "inductive step".

Proposition 4.5. *For a Stein space X we have*

$$\mathrm{Sp}_4^0(\mathcal{O}(X)) = \mathrm{Ep}_4(\mathcal{O}(X)) \cdot \psi(\mathrm{Sp}_2^0(\mathcal{O}(X))).$$

In a similar way one can deduce from our earlier results (Proposition 2.8 and the proof of Theorem 2.3 in [Ivarsson and Kutzschebauch 2012]) the corresponding statements for the special linear groups. The definition of the elementary group E_n and the inclusion ψ of SL_{n-1} into SL_n are the usual ones.

Proposition 4.6. *For a Stein space X and any $n \geq 2$ the group $\mathrm{E}_n(\mathcal{O}(X))$ acts transitively on the set of null-homotopic unimodular rows $U_n^0(\mathcal{O}(X))$.*

Proposition 4.7. *For a Stein space X and any $n \geq 2$ we have*

$$\mathrm{SL}_n^0(\mathcal{O}(X)) = \mathrm{E}_n(\mathcal{O}(X)) \cdot \psi(\mathrm{SL}_{n-1}^0(\mathcal{O}(X))).$$

5. Stratified sprays

We will introduce the concept of a spray associated with a holomorphic submersion following [Gromov 1989; Forstnerič and Prezelj 2002]. First we introduce some notation and terminology. Let $h : Z \to X$ be a holomorphic submersion of a complex manifold Z onto a complex manifold X. For any $x \in X$ the fiber over x of this submersion will be denoted by Z_x. At each point $z \in Z$ the tangent space $T_z Z$ contains *the vertical tangent space* $VT_z Z = \ker Dh$. For holomorphic vector bundles $p : E \to Z$ we denote the zero element in the fiber E_z by 0_z.

Definition 5.1. Let $h : Z \to X$ be a holomorphic submersion of a complex manifold Z onto a complex manifold X. A spray on Z associated with h is a triple (E, p, s), where $p : E \to Z$ is a holomorphic vector bundle and $s : E \to Z$ is a holomorphic map such that for each $z \in Z$ we have

(i) $s(E_z) \subset Z_{h(z)}$,

(ii) $s(0_z) = z$, and

(iii) the derivative $Ds(0_z) : T_{0_z} E \to T_z Z$ maps the subspace $E_z \subset T_{0_z} E$ surjectively onto the vertical tangent space $VT_z Z$.

Remark 5.2. We will also say that the submersion admits a spray. A spray associated with a holomorphic submersion is sometimes called a (fiber-)dominating spray.

One way of constructing dominating sprays, as pointed out by Gromov, is to find finitely many $\mathbb{C}$-complete vector fields that are tangent to the fibers and span the tangent space of the fibers at all points in Z. One can then use the flows φ_j^t of these vector fields V_j to define $s : Z \times \mathbb{C}^N \to Z$ via $s(z, t_1, \dots, t_N) = \varphi_1^{t_1} \circ \cdots \circ \varphi_N^{t_N}(z)$, which gives a spray.

Definition 5.3. Let X and Z be complex spaces. A holomorphic map $h : Z \to X$ is said to be a submersion if for each point $z_0 \in Z$ it is locally equivalent via a fiber-preserving biholomorphic map to a projection $p : U \times V \to U$, where $U \subset X$ is an open set containing $h(z_0)$ and V is an open set in some $\mathbb{C}^d$.

We will need to use stratified sprays, which are defined as follows.

Definition 5.4. We say that a submersion $h : Z \to X$ admits stratified sprays if there is a descending chain of closed complex subspaces $X = X_m \supset \cdots \supset X_0$ such that each stratum $Y_k = X_k \setminus X_{k-1}$ is regular and the restricted submersion $h : Z|_{Y_k} \to Y_k$ admits a spray over a small neighborhood of any point $x \in Y_k$.

Remark 5.5. We say that the stratification $X = X_m \supset \cdots \supset X_0$ is associated with the stratified spray.

In [Forstnerič and Prezelj 2001], see also [Forstnerič 2010, Theorem 8.3], the following is proved.

Theorem 5.6. *Let X be a Stein space with a descending chain of closed complex subspaces $X = X_m \supset \cdots \supset X_0$ such that each stratum $Y_k = X_k \setminus X_{k-1}$ is regular. Assume that $h : Z \to X$ is a holomorphic submersion which admits stratified sprays. Then any continuous section $f_0 : X \to Z$ such that $f_0|_{X_0}$ is holomorphic can be deformed to a holomorphic section $f_1 : X \to Z$ by a homotopy that is fixed on X_0.*

6. Proofs of Lemmas 3.3 and 3.4

Lemmas 3.3 and 3.4 hold for square matrices of any size. In this section we therefore look at $2n \times 2n$ matrices. Given two vectors $\vec{a}$ and $\vec{b}$ in $\mathbb{C}^n$ (i.e., $n \times 1$ matrices), we denote by

$$\begin{pmatrix} \vec{a} \\ \vec{b} \end{pmatrix}$$

the obvious vector in $\mathbb{C}^{2n}$.

We shall consider products of $2n \times 2n$ matrices

$$\begin{pmatrix} I_n & 0 \\ Z_1 & I_n \end{pmatrix} \begin{pmatrix} I_n & W_1 \\ 0 & I_n \end{pmatrix} \begin{pmatrix} I_n & 0 \\ Z_2 & I_n \end{pmatrix} \begin{pmatrix} I_n & W_2 \\ 0 & I_n \end{pmatrix} \cdots,$$

where $Z_1, Z_2, \dots$ and $W_1, W_2, \dots$ are $n \times n$ matrices of variables

$$Z_k = (z_{k,ij}), \quad W_k = (w_{k,ij}), \quad 1 \le i, j \le n.$$

They are symmetric, i.e., $z_{k,ij} = z_{k,ji}$ and $w_{k,ij} = w_{k,ji}$. We call the variables $z_{k,n1}, \dots, z_{k,nn}$ *last row variables* (this term does not apply to the w-variables). If we have K factors, there are $Kn(n+1)/2$ variables. We will also think of the K-tuple $(Z_1, W_1, Z_2, W_2, \dots)$ as a point in $\mathbb{C}^{Kn(n+1)/2}$. We will study the last row of this product, which is a map $\Phi_K : \mathbb{C}^{Kn(n+1)/2} \to \mathbb{C}^{2n} \setminus \{0\}$. We prefer to work with the transpose of this row, which we denote by P^K, a vector in $\mathbb{C}^{2n}$. It follows that

$$P^1 = \begin{pmatrix} \vec{z} \\ \vec{e_n} \end{pmatrix},$$

where $\vec{z} = (z_{1,n1}, \dots, z_{1,nn})^T$ and $\vec{e_n}$ is the last standard basis vector of $\mathbb{C}^n$.

The set S_K for $K \geq 2$ is now defined as the set of K-tuples of symmetric matrices $(Z_1, W_1, \ldots)$ such that in the first $K-1$ matrices all the last row variables (of the Z's) are 0 and the set of all columns of the W's does not span $\mathbb{C}^n$. (This means that the augmented matrix $W_1|W_2|\cdots$ has rank less than n.)

Lemma 6.1. *$P^K : \mathbb{C}^{Kn(n+1)/2} \setminus S_K \to \mathbb{C}^{2n} \setminus \{0\}$ is surjective for $K \geq 3$.*

Proof. We prove the result for $K = 3$. For $K > 3$, simply put $W_2 = Z_3 = W_3 = \cdots = 0$. The proof uses an easy fact from linear algebra; given two vectors $\vec{c}$ and $\vec{d}$ in $\mathbb{C}^n$ with $\vec{c} \neq \vec{0}$ there is a symmetric matrix M such that $M\vec{c} = \vec{d}$. Now let

$$\begin{pmatrix} \vec{a} \\ \vec{b} \end{pmatrix} \in \mathbb{C}^{2n} \setminus \{0\}.$$

Pick any symmetric matrix Z_2 such that $\vec{z} = \vec{a} - Z_2\vec{b} \neq \vec{0}$ and let Z_1 be any symmetric matrix whose last row is $\vec{z}$ and W_1 a symmetric matrix such that $W_1\vec{z} = \vec{b} - \vec{e_n}$. Then $(Z_1, W_1, Z_2) \notin S_3$ and for this choice we have

$$P^3 = \begin{pmatrix} I_n & Z_2 \\ 0 & I_n \end{pmatrix} \begin{pmatrix} I_n & 0 \\ W_1 & I_n \end{pmatrix} \begin{pmatrix} \vec{z} \\ \vec{e_n} \end{pmatrix} = \begin{pmatrix} I_n & Z_2 \\ 0 & I_n \end{pmatrix} \begin{pmatrix} \vec{z} \\ \vec{b} \end{pmatrix} = \begin{pmatrix} \vec{a} \\ \vec{b} \end{pmatrix}. \qquad \square$$

Slightly abusing notation, we denote the Jacobian matrix of Φ_K by JP^K. This is a $(2n \times Kn(n+1)/2)$-matrix whose columns are the derivatives of P^K with respect to one particular variable. We denote the components of P^K by P_i^K, $1 \leq i \leq 2n$. It follows that

$$P^{2k+1} = \begin{pmatrix} I_n & Z_{k+1} \\ 0 & I_n \end{pmatrix} P^{2k}, \tag{6.0.1}$$

$$P^{2k+2} = \begin{pmatrix} I_n & 0 \\ W_{k+1} & I_n \end{pmatrix} P^{2k+1}. \tag{6.0.2}$$

We shall look at the final part of JP^{2k+1}, the part where we differentiate with respect to the new variables $z_{k+1,11}, \ldots, z_{k+1,n1}, z_{k+1,22}, \ldots, z_{k+1,n2}, \ldots, z_{k+1,nn}$. This is a $(2n \times n(n+1)/2)$-matrix. The column where we differentiate with respect to $z_{k+1,ij}$ will consist of P_{n+i}^{2k} in row number j and P_{n+j}^{2k} in row number i. Hence the bottom half of this matrix is zero and we only look at the upper half, an $(n \times n(n+1)/2)$-matrix which we denote by A_{k+1}. If we consider just the columns which contain one particular P_{n+i}^{2k}, we get a square $n \times n$-matrix whose i-th row is $(P_{n+1}^{2k}, \ldots, P_{2n}^{2k})$, has P_{n+i}^{2k} along the diagonal and is otherwise zero. The determinant of this submatrix is $(P_{n+i}^{2k})^n$.

The situation is similar for the final part of JP^{2k+2}, except now the top half is zero and the bottom half B_{k+1} contains $P_1^{2k+1}, \ldots, P_n^{2k+1}$ in the same pattern as for A_{k+1}.

In the proof of the next lemma it will be convenient to use the following notation: if A and B are two matrices with the same column length, we let $A \,|\, B$ denote the matrix obtained by augmenting A with B to the right. By e_{2n} we denote the last vector in the standard basis of $\mathbb{C}^{2n}$.

Lemma 6.2. *P^K is a submersion exactly on the set $\mathbb{C}^{Kn(n+1)/2} \setminus S_K$. If $K = 2k$ and all the last row variables are zero, then $P^{2k} = e_{2n}$ and the span of the bottom half of the JP^{2k} columns equals the span of the columns of $W_1, W_2, \ldots, W_k$.*

Proof. For $N = 1$ the theorem is empty. $P^1 = (z_{1,n1}, \dots, z_{1,nn}, 0, \dots, 0, 1)$ and

$$JP^1 = \begin{pmatrix} I_n \\ 0 \end{pmatrix},$$

where we have removed all zero columns. For $N = 2$ we have

$$P^2 = \begin{pmatrix} I_n & 0 \\ W_1 & I_n \end{pmatrix} P^1.$$

This implies

$$JP^2 = \begin{pmatrix} I_n & 0 \\ W_1 & I_n \end{pmatrix} \begin{pmatrix} I_n \\ 0 \end{pmatrix} \Bigg| \begin{pmatrix} 0 \\ B_1 \end{pmatrix} = \begin{pmatrix} I_n \\ W_1 \end{pmatrix} \Bigg| \begin{pmatrix} 0 \\ B_1 \end{pmatrix},$$

which has full rank if and only if B_1 has full rank. Since $P_i^1 = z_{1,ni}$, by the discussion preceding the lemma, B_1 has full rank if and only if at least one $z_{1,ni}$ is nonzero.

If all $z_{1,ni}$ are zero, then $P^1 = e_{2n}$ and $B_1 = 0$. Hence the statement about the span is trivially true.

We now assume that the theorem is true for $N = 2k$. We have

$$JP^{2k+1} = \begin{pmatrix} I_n & Z_{k+1} \\ 0 & I_n \end{pmatrix} JP^{2k} \Bigg| \begin{pmatrix} A_{k+1} \\ 0 \end{pmatrix}. \tag{6.0.3}$$

If at least one of the previous last row variables is nonzero, then JP^{2k} has full rank by the induction hypothesis and so does JP^{2k+1}. If not, then $P^{2k} = e_{2n}$ and $A_{k+1} = I_n$, after removing zero columns. If $JP^{2k} = \binom{A}{B}$, then

$$JP^{2k+1} = \begin{pmatrix} A + Z_{k+1}B & I_n \\ B & 0 \end{pmatrix},$$

which has full rank if and only if B has full rank. But the column span of B equals the column span of $W_1, \dots, W_k$. This proves the first part of the lemma for $N = 2k + 1$.

If all the previous last row variables are zero, it also follows that

$$P^{2k+1} = (z_{k+1,n1}, \dots, z_{k+1,nn}, 0, \dots, 0, 1)^t.$$

Finally

$$JP^{2k+2} = \begin{pmatrix} I_n & 0 \\ W_{k+1} & I_n \end{pmatrix} JP^{2k+1} \Bigg| \begin{pmatrix} 0 \\ B_{k+1} \end{pmatrix}, \tag{6.0.4}$$

which has full rank if JP^{2k+1} does.

If not, then by the above all the previous last row variables are zero and

$$JP^{2k+2} = \begin{pmatrix} A + Z_{k+1}B & I_n \\ B + W_{k+1}(A + Z_{k+1}B) & W_{k+1} \end{pmatrix} JP^{2k+1} \Bigg| \begin{pmatrix} 0 \\ B_{k+1} \end{pmatrix},$$

which has full rank if and only if at least one $z_{k+1,ni}$ is nonzero by the discussion preceding the lemma. This proves the first part of the lemma for $N = 2k + 2$.

If all the $z_{k+1,ni}$ also are zero, then $P^{2k+1} = e_{2n}$ and so $P^{2k+2} = e_{2n}$. Also $B_{k+1} = 0$ and since the columns of $W_{k+1}(A + Z_{k+1}B)$ are linear combinations of the columns of W_{k+1}, the span of the bottom half of JP^{2k+2} equals the span of the columns of $W_1, \dots, W_{k+1}$ by the induction hypothesis. □

7. The stratification

The goal in this section is to describe the stratification needed to understand that the submersion $\pi_4 \circ \Psi_K : (\mathbb{C}^3)^K \setminus S_K \to \mathbb{C}^4 \setminus \{0\}$ is a stratified elliptic submersion. Let

$$\vec{Z}_K = \begin{cases} (z_1, z_2, z_3, w_1, w_2, w_3, \ldots, w_{3k-2}, w_{3k-1}, w_{3k}) & \text{if } K = 2k, \\ (z_1, z_2, z_3, w_1, w_2, w_3, \ldots, z_{3k+1}, z_{3k+2}, z_{3k+3}) & \text{if } K = 2k+1 \end{cases}$$

and

$$\pi_4 \circ \Psi_K(\vec{Z}_K) = (P_1^K(\vec{Z}_K), P_2^K(\vec{Z}_K), P_3^K(\vec{Z}_K), P_4^K(\vec{Z}_K)).$$

Remark 7.1. We will abuse notation in the following way in the paper. A polynomial not containing a variable can be interpreted as a polynomial of that variable. More precisely, let $L < K$. We have the projection $\pi : \mathbb{C}^K \to \mathbb{C}^L$, $\pi(x_1, \ldots, x_L, \ldots, x_K) = (x_1, \ldots, x_L)$ and $\pi^* : \mathbb{C}[\mathbb{C}^L] \to \mathbb{C}[\mathbb{C}^K]$. For $p \in \mathbb{C}[\mathbb{C}^L]$ we still write p instead of $\pi^*(p)$.

We want to study the fibers

$$\mathcal{F}^K_{(a_1,a_2,a_3,a_4)} = (\pi_4 \circ \Psi_K)^{-1}(a_1, a_2, a_3, a_4).$$

Assume first that $K = 2k+1 \geq 3$ is odd. We see that

$$\pi_4 \circ \Psi_K(\vec{Z}_K) = \pi_4 \circ \Psi_{K-1}(\vec{Z}_{K-1}) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ z_{3k+1} & z_{3k+2} & 1 & 0 \\ z_{3k+2} & z_{3k+3} & 0 & 1 \end{pmatrix}$$

and we get

$$\begin{aligned} P_1^K(\vec{Z}_K) &= P_1^{K-1}(\vec{Z}_{K-1}) + z_{3k+1} P_3^{K-1}(\vec{Z}_{K-1}) + z_{3k+2} P_4^{K-1}(\vec{Z}_{K-1}), \\ P_2^K(\vec{Z}_K) &= P_2^{K-1}(\vec{Z}_{K-1}) + z_{3k+2} P_3^{K-1}(\vec{Z}_{K-1}) + z_{3k+3} P_4^{K-1}(\vec{Z}_{K-1}), \\ P_3^K(\vec{Z}_K) &= P_3^{K-1}(\vec{Z}_{K-1}), \\ P_4^K(\vec{Z}_K) &= P_4^{K-1}(\vec{Z}_{K-1}). \end{aligned}$$

We are led to the equations

$$\begin{aligned} a_1 = P_1^K(\vec{Z}_K) &= P_1^{K-1}(\vec{Z}_{K-1}) + z_{3k+1} P_3^{K-1}(\vec{Z}_{K-1}) + z_{3k+2} P_4^{K-1}(\vec{Z}_{K-1}), \\ a_2 = P_2^K(\vec{Z}_K) &= P_2^{K-1}(\vec{Z}_{K-1}) + z_{3k+2} P_3^{K-1}(\vec{Z}_{K-1}) + z_{3k+3} P_4^{K-1}(\vec{Z}_{K-1}), \\ a_3 = P_3^K(\vec{Z}_K) &= P_3^{K-1}(\vec{Z}_{K-1}), \\ a_4 = P_4^K(\vec{Z}_K) &= P_4^{K-1}(\vec{Z}_{K-1}). \end{aligned} \tag{7.0.1}$$

Notice that these equations simplify to

$$\begin{aligned} a_1 &= P_1^{K-1}(\vec{Z}_{K-1}) + a_3 z_{3k+1} + a_4 z_{3k+2}, \\ a_2 &= P_2^{K-1}(\vec{Z}_{K-1}) + a_3 z_{3k+2} + a_4 z_{3k+3}, \\ a_3 &= P_3^{K-1}(\vec{Z}_{K-1}), \\ a_4 &= P_4^{K-1}(\vec{Z}_{K-1}). \end{aligned}$$

If $(a_3, a_4) \neq (0, 0)$ then we can solve the two first equations for two of the three variables z_{3k+1}, z_{3k+2}, z_{3k+3} and we see that the fiber is a graph over $\mathcal{G}^{K-1}_{(a_3,a_4)} \times \mathbb{C}$, where

$$\mathcal{G}^{K-1}_{(a_3,a_4)} = \{\vec{Z}_{K-1} \in \mathbb{C}^{3K-3} : a_3 = P_3^{K-1}(\vec{Z}_{K-1}),\ a_4 = P_4^{K-1}(\vec{Z}_{K-1})\}.$$

If $(a_3, a_4) = (0, 0)$, we get $\mathcal{F}^K_{(a_1,a_2,0,0)} = \mathcal{F}^{K-1}_{(a_1,a_2,0,0)} \times \mathbb{C}^3$. We see that we get two main cases, namely $(a_3, a_4) = (0, 0)$ and $(a_3, a_4) \neq (0, 0)$. The last case will break into the two subcases, namely $(a_3, a_4) \neq (0, 1)$ and $(a_3, a_4) = (0, 1)$. We need these subcases because $\mathcal{G}^{K-1}_{(0,1)}$ is not smooth. We list the strata below:

- The strata of generic fibers: When $(a_3, a_4) \neq (0, 0)$, the fibers are graphs over $\mathcal{G}^{K-1}_{(a_3,a_4)} \times \mathbb{C}$. This set is divided into two strata as follows:
  - Smooth generic fibers: When $(a_3, a_4) \neq (0, 1)$, the fibers are smooth.
  - Singular generic fibers: When $(a_3, a_4) = (0, 1)$, the fibers are nonsmooth.
- The stratum of nongeneric fibers: When $(a_3, a_4) = (0, 0)$, the fibers are $\mathcal{F}^K_{(a_1,a_2,0,0)} = \mathcal{F}^{K-1}_{(a_1,a_2,0,0)} \times \mathbb{C}^3$. Moreover the fibers are smooth.

We now analyze the case when $K = 2k \geq 3$ is even. Now we have

$$\pi_4 \circ \Psi_K(\vec{Z}_K) = \pi_4 \circ \Psi_{K-1}(\vec{Z}_{K-1}) \begin{pmatrix} 1 & 0 & w_{3k-2} & w_{3k-1} \\ 0 & 1 & w_{3k-1} & w_{3k} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and $\mathcal{F}^K_{(a_1,a_2,a_3,a_4)}$ is the solution set of the equations

$$\begin{aligned} a_1 &= P_1^K(\vec{Z}_K) = P_1^{K-1}(\vec{Z}_{K-1}), \\ a_2 &= P_2^K(\vec{Z}_K) = P_2^{K-1}(\vec{Z}_{K-1}), \\ a_3 &= P_3^K(\vec{Z}_K) = P_3^{K-1}(\vec{Z}_{K-1}) + w_{3k-2} P_1^{K-1}(\vec{Z}_{K-1}) + w_{3k-1} P_2^{K-1}(\vec{Z}_{K-1}), \\ a_4 &= P_4^K(\vec{Z}_K) = P_4^{K-1}(\vec{Z}_{K-1}) + w_{3k-1} P_1^{K-1}(\vec{Z}_{K-1}) + w_{3k} P_2^{K-1}(\vec{Z}_{K-1}). \end{aligned} \tag{7.0.2}$$

As in the previous case these equations simplify:

$$\begin{aligned} a_1 &= P_1^{K-1}(\vec{Z}_{K-1}), \\ a_2 &= P_2^{K-1}(\vec{Z}_{K-1}), \\ a_3 &= P_3^{K-1}(\vec{Z}_{K-1}) + a_1 w_{3k-2} + a_2 w_{3k-1}, \\ a_4 &= P_4^{K-1}(\vec{Z}_{K-1}) + a_1 w_{3k-1} + a_2 w_{3k}. \end{aligned}$$

Let

$$\mathcal{H}^{K-1}_{(a_1,a_2)} = \{\vec{Z}_{K-1} \in \mathbb{C}^{3K-3} : a_1 = P_1^{K-1}(\vec{Z}_{K-1}),\ a_2 = P_2^{K-1}(\vec{Z}_{K-1})\}.$$

An analysis similar to that above gives us the following strata:

- The stratum of generic fibers: When $(a_1, a_2) \neq (0, 0)$, the fibers are graphs over $\mathcal{H}^{K-1}_{(a_1,a_2)} \times \mathbb{C}$. Moreover the fibers are smooth.

• The strata of nongeneric fibers: When $(a_1, a_2) = (0, 0)$, the fibers are $\mathcal{F}^K_{(0,0,a_3,a_4)} = \mathcal{F}^{K-1}_{(0,0,a_3,a_4)} \times \mathbb{C}^3$. This set is divided into two strata as follows:

- Smooth nongeneric fibers: When $(a_3, a_4) \neq (0, 1)$, the fibers are smooth.
- Singular nongeneric fibers: When $(a_3, a_4) = (0, 1)$, the fibers are nonsmooth.

8. Determination of complete vector fields

The description of the fibers in Section 7 leads us to study vector fields simultaneously tangent to the level sets $\{P = c_1\}$, $\{Q = c_2\}$ of two functions $P, Q : \mathbb{C}^N \to \mathbb{C}$. Such fields can be constructed in the following way. Pick three variables x, y, z from the variables $x_1, \dots, x_N$ on $\mathbb{C}^N$ and consider the vector fields

$$D_{xyz}(P, Q) = \det \begin{pmatrix} \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \partial P/\partial x & \partial P/\partial y & \partial P/\partial z \\ \partial Q/\partial x & \partial Q/\partial y & \partial Q/\partial z \end{pmatrix}, \tag{8.0.1}$$

which are simultaneously tangent to the level sets. As mentioned in Section 5 we want to use a finite collection of complete vector fields spanning tangent space at every point to prove (stratified) ellipticity. It is an easy exercise to show that the collection of these vector fields over all possible triples spans the tangent space at smooth points of the variety $\{P = c_1\} \cap \{Q = c_2\}$. It turns out that many of the vector fields we get by this method are complete but unfortunately not all of them. The complete vector fields from this collection will not span the tangent space at all points for all level sets. To overcome this difficulty and still producing dominating sprays from this collection of available complete fields is the main technical part of our paper explained in Section 9.

Now we will begin to describe the complete vector fields tangent to the fibers of $\pi_4 \circ \Psi_K = (P_1^K, P_2^K, P_3^K, P_4^K)$ that we get using (8.0.1). It will be convenient to group the variables as in Section 6, $Z_1, W_1, Z_2, W_2, \dots$, where

$$Z_k = \begin{pmatrix} z_{3k-2} & z_{3k-1} \\ z_{3k-1} & z_{3k} \end{pmatrix}$$

and similarly for W_k. Since the variable z_1 never enters in P^K, we omit it from the first group Z_1. Note that $P^1 = (z_2, z_3, 0, 1)^T$. We are going to study the vector fields

$$V_{ij}^K(x, y, z) = D_{xyz}(P_i^K, P_j^K).$$

The 2×2 minors occurring as coefficients are denoted by $C_{ij}^K(\cdot, \cdot)$, i.e.,

$$V_{ij}^K(x, y, z) = C_{ij}^K(y, z)\frac{\partial}{\partial x} - C_{ij}^K(x, z)\frac{\partial}{\partial y} + C_{ij}^K(x, y)\frac{\partial}{\partial z}.$$

The description of the complete vector fields will be done inductively. We start with $K = 2$. We have to study $\mathcal{G}^2_{(a_3,a_4)}$, or equivalently, the equations

$$\begin{aligned} a_3 &= P_3^2(z_1, \dots, w_3) = z_2 w_1 + z_3 w_2, \\ a_4 &= P_4^2(z_1, \dots, w_3) = 1 + z_2 w_2 + z_3 w_3. \end{aligned} \tag{8.0.2}$$

We are interested in which triples (x, y, z) of variables from the list z_2, z_3, w_1, w_2, w_3 give complete vector fields $V_{34}^2(x, y, z)$ and we denote the set of these triples by $\mathcal{T}_2$. By definition $\mathcal{T}_1 = \varnothing$.

An easy computation gives that

$$\mathcal{T}_2 = \{(w_1, w_2, w_3), (z_2, w_2, w_3), (z_3, w_1, w_2), \\ (z_2, w_1, w_3), (z_3, w_1, w_3), (z_2, z_3, w_1), (z_2, z_3, w_3)\}. \tag{8.0.3}$$

For all the remaining noncomplete triples there is a variable such that the equation is quadratic for that variable. We are now interested in determining at every stage the triples of variables (x, y, z) such that $V_{12}^{2k+1}(x, y, z)$, for $K = 2k + 1$ odd, and $V_{34}^{2k+2}(x, y, z)$, for $K = 2k + 2$ even, are complete. We shall denote the set of such triples by $\mathcal{T}_K$. The terms occurring in P^K are of degree 1 in the occurring variables; hence the coefficients C_{ij}^K are either of degree 1 or 2 in the occurring variables. A triple giving a coefficient which is quadratic in the integration variable (for instance if $C_{ij}^K(y, z)$ is quadratic in the x-variable) will not be complete, and we shall refer to such a triple as a quadratic triple and the corresponding vector field as a quadratic vector field. The content of the next lemma is that all the remaining triples give complete vector fields. The variables that do not occur in a triple will have constant solutions and are therefore treated as such in the proof.

Lemma 8.1. *For $k \geq 1$, we have $\mathcal{T}_{2k} \subset \mathcal{T}_{2k+1} \subset \mathcal{T}_{2k+2}$. Moreover*

$$\begin{aligned} \mathcal{T}_{2k+1} \setminus \mathcal{T}_{2k} = {} & \{(z_{3k+1}, z_{3k+2}, z_{3k+3})\} \\ & \cup \{(w_{3k-2}, z_{3k+1}, z_{3k+3}), (w_{3k-2}, z_{3k+2}, z_{3k+3})\} \\ & \cup \{(w_{3k}, z_{3k+1}, z_{3k+2}), (w_{3k}, z_{3k+1}, z_{3k+3})\} \\ & \cup \{(a, b, z_{3k+1}), (a, b, z_{3k+3}) : \textit{a and b are from the same group}\} \\ & \cup \{(a, b, z_{3k+1}) : \textit{a the last variable of one group and b the first of the next}\} \\ & \cup \{(a, b, z_{3k+3}) : \textit{a the last variable of one group and b the first of the next}\} \\ & \cup \{(a, b, z_{3k+1}) : \textit{a the first variable of one group and b the last of the next}\} \\ & \cup \{(a, b, z_{3k+3}) : \textit{a the first variable of one group and b the last of the next}\} \end{aligned} \tag{8.0.4}$$

and

$$\begin{aligned} \mathcal{T}_{2k+2} \setminus \mathcal{T}_{2k+1} = {} & \{(w_{3k+1}, w_{3k+2}, w_{3k+3})\} \\ & \cup \{(z_{3k+1}, w_{3k+1}, w_{3k+3}), (z_{3k+1}, w_{3k+2}, w_{3k+3})\} \\ & \cup \{(z_{3k+3}, w_{3k+1}, w_{3k+2}), (z_{3k+3}, w_{3k+1}, w_{3k+3})\} \\ & \cup \{(a, b, w_{3k+1}), (a, b, w_{3k+3}) : \textit{a and b are from the same group}\} \\ & \cup \{(a, b, w_{3k+1}) : \textit{a the last variable of one group and b the first of the next}\} \\ & \cup \{(a, b, w_{3k+3}) : \textit{a the last variable of one group and b the first of the next}\} \\ & \cup \{(a, b, w_{3k+1}) : \textit{a the first variable of one group and b the last of the next}\} \\ & \cup \{(a, b, w_{3k+3}) : \textit{a the first variable of one group and b the last of the next}\}. \end{aligned} \tag{8.0.5}$$

In combination with (8.0.3) this gives us a complete description of the sets $\mathcal{T}_L$, $L \geq 2$.

Proof. The result is true for $\mathcal{T}_2$. The first group is interpreted as $\{z_2, z_3\}$ and z_{3k+1} must be replaced by z_2. The missing triplets are precisely the quadratic triples.

We shall prove (8.0.4), the proof of (8.0.5) being identical. There is a lot of symmetry in the proof and we will not repeat arguments already given in a situation symmetric to a proven statement. We first consider triples (x, y, z) not containing any variables from the new group Z_{k+1}, i.e., z_{3k+1}, z_{3k+2} and z_{3k+3}. It then follows from (6.0.1) (omitting variables for shorter notation) that

$$V_{12}^{2k+1} = V_{12}^{2k} + z_{3k+1} V_{32}^{2k} - z_{3k+2} V_{24}^{2k} - z_{3k+2} V_{31}^{2k} + z_{3k+3} V_{14}^{2k} + (z_{3k+1} z_{3k+3} - z_{3k+2}^2) V_{34}^{2k}. \tag{8.0.6}$$

A quadratic triple will still be quadratic since V_{34}^{2k} is. For a triple in $\mathcal{T}_{2k}$, notice that in all of the first five terms the V_{ij}^{2k} is obtained by replacing one or two of the functions P_3^{2k} and P_4^{2k} by P_1^{2k} and/or P_2^{2k}. By (6.0.2) all of the terms occurring in P_1^{2k} or P_2^{2k} divide a term occurring in P_3^{2k} and also a term occurring in P_4^{2k}. This means that all terms occurring in the first five vector fields above are already present in V_{34}^{2k} and completeness is not destroyed. We also notice that for any pair x, y of previous variables, the coefficient $C_{12}^{2k+1}(x, y)$ will also satisfy (8.0.6).

We next consider triples containing some of the new variables z_{3k+1}, z_{3k+2} and z_{3k+3}. The Jacobian matrix is now given by (6.0.3), where

$$A^{k+1} = \begin{pmatrix} P_3^{2k} & P_4^{2k} & 0 \\ 0 & P_3^{2k} & P_4^{2k} \end{pmatrix}. \tag{8.0.7}$$

If the triple contains all three variables, then

$$V_{12}^{2k+1}(z_{k+1}, z_{k+2}, z_{k+3}) = (P_4^{2k})^2 \frac{\partial}{\partial z_{3k+1}} - (P_3^{2k})(P_4^{2k}) \frac{\partial}{\partial z_{3k+2}} + (P_3^{2k})^2 \frac{\partial}{\partial z_{3k+3}},$$

and the coefficients do not contain any of the Z_{k+1}-variables; hence this is complete. (The solutions are just affine functions.)

We now consider the case of two new variables. The first possibility is (x, z_{3k+1}, z_{3k+2}). The coefficient of $\partial/\partial x$ is $(P_3^{2k})^2$. Since P_3^{2k} contains all previous variables except w_{3k}, this is quadratic in all those variables and $x = w_{3k}$ is the only possibility. The solution for w_{3k} is affine. The coefficient of $\partial/\partial z_{3k+2}$ is now

$$-\left(\frac{\partial P_2^{2k}}{\partial w_{3k}} + z_{k+3} \frac{\partial P_4^{2k}}{\partial w_{3k}} \right),$$

which is just a constant and the solution is again affine. Finally the coefficient of $\partial/\partial z_{3k+1}$ is given by

$$\frac{\partial P_1^{2k}}{\partial w_{3k}} + z_{k+2} \frac{\partial P_4^{2k}}{\partial w_{3k}},$$

which is an affine function and the solution is entire. Hence this field is complete.

Precisely the same logic applies to the triple (x, z_{3k+2}, z_{3k+3}) except now w_{3k-2} is the only missing variable (now in P_4^{2k}).

The final possibility of two new variables is the triple (x, z_{3k+1}, z_{3k+3}). The coefficient of $\partial/\partial x$ is now $P_3^{2k} P_4^{2k}$ which is of degree 1 in w_{3k-2} and w_{3k} and quadratic in all other previous variables. We consider the case of $x = w_{3k-2}$, the case $x = w_{3k}$ being identical. The coefficient is an affine function of w_{3k-2}; hence the solution is entire. The coefficient of $\partial/\partial z_{3k+1}$ is $-z_{3k+1} P_1^{2k-1} P_4^{2k}$, which is just a linear function of z_{3k+1} and the solution is entire. The coefficient of $\partial/\partial z_{3k+3}$ is $-z_{3k+2}P_2^{2k-1}P_4^{2k}$, which is just a constant and the solution is affine.

We finally consider the case of one new variable and two previous variables x, y. It follows that $C_{12}^{2k+1}(x, y)$ satisfies (8.0.6), and hence is quadratic in z_{3k+2}, so this cannot be the new variable. In order to investigate z_{3k+1} and z_{3k+3} we need to understand which variables are involved in the coefficients. To do this we look at each previous group of variables Z_j and W_j for $1 \le j \le k$ and see which variables are involved in the first two rows of the Jacobian with respect to these variables at level $2k+1$. For a Z_j-group we need to consider the matrix

$$\begin{pmatrix} \partial P_1^{2k+1}/\partial z_{3j-2} & \partial P_1^{2k+1}/\partial z_{3j-1} & \partial P_1^{2k+1}/\partial z_{3j} \\ \partial P_2^{2k+1}/\partial z_{3j-2} & \partial P_2^{2k+1}/\partial z_{3j-1} & \partial P_2^{2k+1}/\partial z_{3j} \end{pmatrix}$$

and the same for a W_j-group. The Z_1-group only consists of z_2 and z_3. The Z_j-variables do not occur in the above matrix. There is a simple formula for the above matrix which follows from (6.0.3) and (6.0.4). The matrix is the first two rows of the matrix $(I = I_2)$:

$$\begin{pmatrix} I & Z_{k+1} \\ 0 & I \end{pmatrix} \cdots \begin{pmatrix} I & 0 \\ W_j & I \end{pmatrix} \begin{pmatrix} A_j \\ 0 \end{pmatrix}$$

and this formula makes it easy to track which variables are missing at each step, in addition to the Z_j-variables. We arrive at the following matrix of missing variables:

$$\begin{pmatrix} w_{3j-3}, w_{3j}, z_{3k+3} & z_{3k+3} & w_{3j-5}, w_{3j-2}, z_{3k+3} \\ w_{3j-3}, w_{3j}, z_{3k+1} & z_{3k+1} & w_{3j-5}, w_{3j-2}, z_{3k+1} \end{pmatrix}.$$

In the case $j = 1$ the missing-variable matrix is

$$\begin{pmatrix} w_3, z_{3k+3} & w_1, z_{3k+3} \\ w_3, z_{3k+1} & w_1, z_{3k+1} \end{pmatrix}.$$

We now consider a W_j-group. Again the W_j-variables do not enter. We now have to consider the first two rows of the matrix

$$\begin{pmatrix} I & Z_{k+1} \\ 0 & I \end{pmatrix} \cdots \begin{pmatrix} I & Z_{j+1} \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 \\ B_j \end{pmatrix},$$

and this leads to the following missing-variable matrix for $j < k$:

$$\begin{pmatrix} z_{3j}, z_{3j+3}, z_{3k+3} & z_{3k+3} & z_{3j-2}, z_{3j+1}, z_{3k+3} \\ z_{3j}, z_{3j+3}, z_{3k+1} & z_{3k+1} & z_{3j-2}, z_{3j+1}, z_{3k+1} \end{pmatrix}.$$

For $j = 1$ we replace z_{3j-2} by z_2. For $j = k$ the middle entries in the upper-left and the lower-right corners are replaced by z_{3k+2}.

We first investigate triples (x, y, z_{3k+1}), where x and y are not from Z_{k+1}. If x and y are from the same group, then since z_{3k+1} occurs in every entry in the second row of the missing-variable matrix, $C_{12}^{2k+1}(x, z_{3k+1})$ and $C_{12}^{2k+1}(y, z_{3k+1})$ do not depend on any of the variables x, y, z_{3k+1}; hence x and y are both affine functions. $C_{12}^{2k+1}(x, y)$ does not depend on x, y and is of degree 1 in z_{3k+1}; hence the solution is entire.

Now assume that x and y are from different groups. If x is not a missing variable in $\partial P_2^{2k+1}/\partial y$, then y is not a missing variable in $\partial P_2^{2k+1}/\partial x$. The variables x and y are not both w_{3k}; let's say x is. Then

$$C_{12}^{2k+1}(y, z_{3k+1}) = -\left(\frac{\partial P_2^{2k+1}}{\partial y}\right) P_3^{2k}$$

is quadratic in x and the field is not complete. Hence x and y must both appear in the second row of the missing-variable matrix of each other.

We now look at possibilities for x and y. Assume first that x is in Z_j group with $1 < j \le k$. There are now four possibilities:

- $x = z_{3j-2}$ in which case $y = w_{3j-3}$ or $y = w_{3j}$, or
- $x = z_{3j}$ in which case $y = w_{3j-5}$ or $y = w_{3j-2}$.

We consider the first case. Then

$$C_{12}^{2k+1}(w_{3j-3}, z_{3j-2}) = \frac{\partial P_1^{2k+1}}{\partial w_{3j-3}} \frac{\partial P_2^{2k+1}}{\partial z_{3j-2}} - \frac{\partial P_2^{2k+1}}{\partial w_{3j-3}} \frac{\partial P_1^{2k+1}}{\partial z_{3j-2}},$$

and from the missing-variable matrix we see that this does not depend on z_{3j-2} and w_{3j-3} and is of degree 1 in z_{3k+1}; hence we have an entire solution for z_{3k+1}. We also have

$$C_{12}^{2k+1}(w_{3j-3}, z_{3k+1}) = -\frac{\partial P_2^{2k+1}}{\partial w_{3j-3}} P_3^{2k},$$

$$C_{12}^{2k+1}(z_{3j-2}, z_{3k+1}) = -\frac{\partial P_2^{2k+1}}{\partial z_{3j-2}} P_3^{2k}.$$

The partial derivatives on the right-hand sides do not depend on any of the variables in the triple, and hence are just constants. It also follows from the missing-variable matrix that P_3^{2k} does not contain the product of z_{3j-2} and w_{3j-3}; hence the equations for these two variables form a linear system with constant coefficients. This has an entire solution. The three other cases all have similar structure and have entire solutions. In the case $j = 1$, we either have $x = z_2$ and $y = w_3$, or $x = z_3$ and $y = w_1$ and the discussion is the same. It also follows from the missing-variable matrix that x and y cannot come from different W-groups. This proves the result in the case of picking z_{3k+1} from the last group. The proof in the case of picking z_{3k+3} from the last group is completely symmetric. This provides the final detail in the proof. □

In order to produce complete fields that are also tangential to fibers of the submersion, we introduce the following notation and terminology.

Definition 8.2. Let $\Xi_3 = \mathcal{T}_2$. For $K \ge 4$ let

$$\Xi_K = \mathcal{T}_{K-1} \setminus \mathcal{T}_{K-2}.$$

We say that the triples in Ξ_K are *introduced on level K*.

We will now use these complete fields to produce complete fields which are tangential to the fibers $\mathcal{F}^K_{(a_1,a_2,a_3,a_4)}$. Here we will use triples introduced on level K to produce complete tangential fields.

First consider the case $K = 2k+1 \geq 3$ odd.

If $a_3 \neq 0$, we use (7.0.1) to get

$$z_{3k+2} = \frac{1}{a_3}(a_2 - P_2^{2k}(\vec{Z}_{2k}) - a_4 z_{3k+3})$$

and

$$\begin{aligned} z_{3k+1} &= \frac{1}{a_3}(a_1 - P_1^{2k}(\vec{Z}_{2k}) - a_4 z_{3k+2}) \\ &= \frac{1}{a_3}\Big(a_1 - P_1^{2k}(\vec{Z}_{2k}) - \frac{a_4}{a_3}(a_2 - P_2^{2k}(\vec{Z}_{2k}) - z_{3k+3}P_4^{2k}(\vec{Z}_{2k}))\Big) \\ &= \frac{1}{a_3^2}(a_1 a_3 - a_3 P_1^{2k}(\vec{Z}_{2k}) - a_2 a_4 + a_4 P_2^{2k}(\vec{Z}_{2k}) + a_4^2 z_{3k+3}). \end{aligned}$$

Using this we define a biholomorphism

$$\alpha : \mathcal{G}^{2k}_{(a_3,a_4)} \times \mathbb{C}_{z_{3k+3}} \to \mathcal{F}^K_{(a_1,a_2,a_3,a_4)}.$$

On

$$\mathcal{G}^{2k}_{(a_3,a_4)} \times \mathbb{C}_{z_{3k+3}}$$

we have the complete fields $\partial^{2k}_{x_1x_2x_3}$ for x_1, x_2, x_3 in Ξ_{2k+1} and also the complete field $\partial/\partial z_{3k+3}$. Using the biholomorphism α we get complete fields on $\mathcal{F}^K_{(a_1,a_2,a_3,a_4)}$ for $a_3 \neq 0$ of the form

$$\theta^{2k+1,*}_{x_1x_2x_3} = \partial^{2k}_{x_1x_2x_3} + \frac{1}{a_3}\partial^{2k}_{x_1x_2x_3}(P_2^{2k}(\vec{Z}_{2k}))\frac{\partial}{\partial z_{3k+2}} + \frac{1}{a_3^2}\partial^{2k}_{x_1x_2x_3}(a_3 P_1^{2k}(\vec{Z}_{2k}) - a_4 P_2^{2k}(\vec{Z}_{2k}))\frac{\partial}{\partial z_{3k+1}} \tag{8.0.8}$$

and

$$\gamma^{2k+1,*} = \frac{\partial}{\partial z_{3k+3}} + \frac{a_4^2}{a_3^2}\frac{\partial}{\partial z_{3k+1}} - \frac{a_4}{a_3}\frac{\partial}{\partial z_{3k+2}}. \tag{8.0.9}$$

Since $P_3^{2k} = a_3$ and $P_4^{2k} = a_4$ on the fiber, we get meromorphic fields on $(\mathbb{C}^3)^K$

$$\begin{aligned} \theta^{2k+1,*}_{x_1x_2x_3} = \partial^{2k}_{x_1x_2x_3} &+ \frac{\partial^{2k}_{x_1x_2x_3}(P_2^{2k}(\vec{Z}_{2k}))}{P_3^{2k}(\vec{Z}_{2k})}\frac{\partial}{\partial z_{3k+2}} \\ &+ \left(\frac{\partial^{2k}_{x_1x_2x_3}(P_1^{2k}(\vec{Z}_{2k}))}{P_3^{2k}(\vec{Z}_{2k})} - \frac{P_4^{2k}(\vec{Z}_{2k})\partial^{2k}_{x_1x_2x_3}(P_2^{2k}(\vec{Z}_{2k}))}{P_3^{2k}(\vec{Z}_{2k})^2}\right)\frac{\partial}{\partial z_{3k+1}} \end{aligned} \tag{8.0.10}$$

and

$$\gamma^{2k+1,*} = \frac{\partial}{\partial z_{3k+3}} + \frac{P_4^{2k}(\vec{Z}_{2k})^2}{P_3^{2k}(\vec{Z}_{2k})^2}\frac{\partial}{\partial z_{3k+1}} - \frac{P_4^{2k}(\vec{Z}_{2k})}{P_3^{2k}(\vec{Z}_{2k})}\frac{\partial}{\partial z_{3k+2}} \tag{8.0.11}$$

(abusing notation), with poles on $P_3^{2k} = 0$. Since P_3^{2k} is in the kernel of these fields, we can multiply the fields by $(P_3^{2k})^2$ and get the following complete fields that are globally defined on $(\mathbb{C}^3)^K$ and preserve the fibers of $\pi_4 \circ \Psi_K$:

$$\begin{aligned} \theta^{2k+1}_{x_1x_2x_3} &= P_3^{2k}(\vec{Z}_{2k})^2\theta^{2k+1,*}_{x_1x_2x_3} \\ &= P_3^{2k}(\vec{Z}_{2k})^2\partial^{2k}_{x_1x_2x_3} + P_3^{2k}(\vec{Z}_{2k})\partial^{2k}_{x_1x_2x_3}(P_2^{2k}(\vec{Z}_{2k}))\frac{\partial}{\partial z_{3k+2}} \\ &\quad + \big[P_3^{2k}(\vec{Z}_{2k})\partial^{2k}_{x_1x_2x_3}(P_1^{2k}(\vec{Z}_{2k})) - P_4^{2k}(\vec{Z}_{2k})\partial^{2k}_{x_1x_2x_3}(P_2^{2k}(\vec{Z}_{2k}))\big]\frac{\partial}{\partial z_{3k+1}} \end{aligned} \tag{8.0.12}$$

for $x_1, x_2, x_3 \in \Xi_{2k+1}$, and

$$\begin{aligned}\gamma^{2k+1} &= P_3^{2k}(\vec{Z}_{2k})^2\gamma^{2k+1,*}\\ &= P_3^{2k}(\vec{Z}_{2k})^2\frac{\partial}{\partial z_{3k+3}} + P_4^{2k}(\vec{Z}_{2k})^2\frac{\partial}{\partial z_{3k+1}} - P_3^{2k}(\vec{Z}_{2k})P_4^{2k}(\vec{Z}_{2k})\frac{\partial}{\partial z_{3k+2}}.\end{aligned}\tag{8.0.13}$$

If $a_4 \neq 0$ we can define a biholomorphism

$$\beta : \mathcal{G}^{2k}_{(a_3,a_4)} \times \mathbb{C}_{z_{3k+1}} \to \mathcal{F}^K_{(a_1,a_2,a_3,a_4)}$$

using (7.0.1) and

$$z_{3k+2} = \frac{1}{a_4}(a_1 - P_1^{2k}(\vec{Z}_{2k}) - a_3 z_{3k+1})$$

and

$$\begin{aligned}z_{3k+3} &= \frac{1}{a_4}(a_2 - P_2^{2k}(\vec{Z}_{2k}) - a_3 z_{3k+2})\\ &= \frac{1}{a_4}\Big(a_2 - P_2^{2k}(\vec{Z}_{2k}) - \frac{a_3}{a_4}(a_1 - P_1^{2k}(\vec{Z}_{2k}) - a_3 z_{3k+1})\Big)\\ &= \frac{1}{a_4^2}(a_2a_4 - a_4P_2^{2k}(\vec{Z}_{2k}) - a_1a_3 + a_3P_1^{2k}(\vec{Z}_{2k}) + a_3^2 z_{3k+1}).\end{aligned}$$

On

$$\mathcal{G}^{2k}_{(a_3,a_4)} \times \mathbb{C}_{z_{3k+1}}$$

we have the complete fields $\partial^{2k}_{x_1x_2x_3}$ for x_1, x_2, x_3 in Ξ_{2k+1} and $\partial/\partial z_{3k+1}$. Proceeding as above we get the complete fields

$$\begin{aligned}\phi^{2k+1}_{x_1x_2x_3} &= P_4^{2k}(\vec{Z}_{2k})^2\partial^{2k}_{x_1x_2x_3} + P_4^{2k}(\vec{Z}_{2k})\partial^{2k}_{x_1x_2x_3}(P_1^{2k}(\vec{Z}_{2k}))\frac{\partial}{\partial z_{3k+2}}\\ &\quad + \big[P_4^{2k}(\vec{Z}_{2k})\partial^{2k}_{x_1x_2x_3}(P_2^{2k}(\vec{Z}_{2k})) - P_3^{2k}(\vec{Z}_{2k})\partial^{2k}_{x_1x_2x_3}(P_1^{2k}(\vec{Z}_{2k}))\big]\frac{\partial}{\partial z_{3k+3}}\end{aligned}\tag{8.0.14}$$

for $x_1, x_2, x_3 \in \Psi_{2k+1}$. The field γ^{2k+1} is the same as in the case $a_3 \neq 0$.

For the case $K = 2k \geq 3$ even, an analogous procedure leads to the following complete fields on $(\mathbb{C}^3)^K$ tangent to the fibers of $\pi_4 \circ \Psi_K$:

$$\begin{aligned}\theta^{2k}_{x_1x_2x_3} &= P_1^{2k-1}(\vec{Z}_{2k-1})^2\partial^{2k-1}_{x_1x_2x_3} + P_1^{2k-1}(\vec{Z}_{2k-1})\partial^{2k-1}_{x_1x_2x_3}(P_4^{2k-1}(\vec{Z}_{2k-1}))\frac{\partial}{\partial w_{3k-1}}\\ &\quad + \big[P_1^{2k-1}(\vec{Z}_{2k-1})\partial^{2k-1}_{x_1x_2x_3}(P_3^{2k-1}(\vec{Z}_{2k-1})) - P_2^{2k-1}(\vec{Z}_{2k-1})\partial^{2k-1}_{x_1x_2x_3}(P_4^{2k-1}(\vec{Z}_{2k-1}))\big]\frac{\partial}{\partial w_{3k-2}}\end{aligned}\tag{8.0.15}$$

for $x_1, x_2, x_3 \in \Xi_{2k}$,

$$\begin{aligned}\phi^{2k}_{x_1x_2x_3} &= P_2^{2k-1}(\vec{Z}_{2k-1})^2\partial^{2k-1}_{x_1x_2x_3} + P_2^{2k-1}(\vec{Z}_{2k-1})\partial^{2k-1}_{x_1x_2x_3}(P_3^{2k-1}(\vec{Z}_{2k-1}))\frac{\partial}{\partial w_{3k-1}}\\ &\quad + \big[P_2^{2k-1}(\vec{Z}_{2k-1})\partial^{2k-1}_{x_1x_2x_3}(P_4^{2k-1}(\vec{Z}_{2k-1})) - P_1^{2k-1}(\vec{Z}_{2k-1})\partial^{2k-1}_{x_1x_2x_3}(P_3^{2k-1}(\vec{Z}_{2k-1}))\big]\frac{\partial}{\partial w_{3k}}\end{aligned}\tag{8.0.16}$$

for $x_1, x_2, x_3 \in \Xi_{2k}$, and

$$\gamma^{2k} = P_1^{2k-1}(\vec{Z}_{2k-1})^2\frac{\partial}{\partial w_{3k}} + P_2^{2k-1}(\vec{Z}_{2k-1})^2\frac{\partial}{\partial w_{3k-2}} - P_1^{2k-1}(\vec{Z}_{2k-1})P_2^{2k-1}(\vec{Z}_{2k-1})\frac{\partial}{\partial w_{3k-1}}.\tag{8.0.17}$$

Remark 8.3. It follows from the inductive formulas (7.0.1) and (7.0.2) that $\theta^K_{x_1x_2x_3}$, $\phi^K_{x_1x_2x_3}$ and γ^K, considered as vector fields on $(\mathbb{C}^3)^L$, are tangent to the fibers $\mathcal{F}^L_{(a_1,a_2,a_3a_4)}$ for $L \geq K$. In other words, the fields associated with triples introduced on level K are tangential to all fibers $\mathcal{F}^L$ for $L \geq K$.

9. Strategy of proof of stratified ellipticity

We outline the strategy for proving that the submersion is a stratified elliptic submersion. We have seen that the fibers are given by four polynomial equations. We have also seen that these four equations can be reduced to two equations. We then use the exact form of these two equations to find Ξ_K so that $\partial^K_{x_1x_2x_3}$ are complete vector fields exactly when $x_1, x_2, x_3 \in \Xi_K$. This leads us to the globally defined complete vector fields $\theta^K_{x_1x_2x_3}$, $\phi^K_{x_1x_2x_3}$ and γ^K described in Section 8. Find a big (a complement of an analytic subset) "good" set on the fibers where the collection of these vector fields spans the tangent space of the fiber. For points outside the good set find a complete field V whose orbit through the point intersects the good set. At points along the orbit that are also in the good set, the collection of complete vector fields above spans. Now pull back the collection of vector fields by suitable flow automorphisms of V and add these fields to the collection (see Definition 10.7). This enlarged collection of complete vector fields spans in a bigger set, thus enlarging the good set. Continue this enlarging of the collection of vector fields until it spans the tangent space at every point of every fiber in the stratum. To accomplish this strategy we need the following technical results.

Lemma 9.1. *Let M be a Stein manifold and $N_0 \subset N \subset M$ analytic subvarieties. Given a finite collection $\theta_1, \dots, \theta_k$ of complete holomorphic vector fields on M which span the tangent space T_xM at all points $x \in M \setminus N$ and given another complete holomorphic vector field ϕ on M (whose flow we denote by $\alpha_t \in \mathrm{Aut_{hol}}(M)$, $t \in \mathbb{C}$) with the property that the orbit through points of $N \setminus N_0$ is leaving N; i.e., for all $x \in N \setminus N_0$ we have $\{\alpha_t(x) : t \in \mathbb{C}\} \not\subset N$. Then there are finitely many times $t_i \in \mathbb{C}$, $i = 1, \dots, l$, such that for all $x \in N \setminus N_0$ we have $\{\alpha_{t_i}(x)\}_{i=1}^{l} \not\subset N$. In particular the finite collection $\{\alpha^\star_{t_i}(\theta_m)\}_{i=1,m=1}^{l,k}$ of complete holomorphic vector fields on M spans the tangent space T_xM at all points $x \in M \setminus N_0$.*

Proof. The analytic subset N has at most countably many components. Denote by B_i those components which are not entirely contained in N_0. Define a_0 to be the maximal dimension of them. Choose a point x_i from each of those B_i. For every i the set $A_i := \{t \in \mathbb{C} : \alpha_t(x_i) \in N\}$ is discrete. Since a countable union of discrete sets is meager in $\mathbb{C}$, we can find t_1 such that $t_1 \notin A_i$ for all i. Denote by $\widetilde{B}_i$ those components of the analytic subset $N_1 := \{y \in N : \alpha_{t_1}(y) \in N\}$ which are not entirely contained in N_0 and define a_1 to be the maximal dimension of them. By construction $a_1 < a_0$. Choose a point $\tilde{x}_i$ from each of those $\widetilde{B}_i$. For every i the set $\tilde{A}_i := \{t \in \mathbb{C} : \alpha_t(\tilde{x}_i) \in N\}$ is discrete. Since a countable union of discrete sets is meager in $\mathbb{C}$, we can find t_2 such that $t_2 \notin \tilde{A}_i$ for all i.

Let a_2 be the maximal dimension of those components of the analytic subset $N_2 := \{y \in N : \alpha_{t_1}(y) \in N \text{ and } \alpha_{t_2}(y) \in N\}$ which are not entirely contained in N_0. By construction $a_2 < a_1$ and continuing the construction after finitely steps we reach our conclusion. □

The next lemma is a generalized and parametrized version of the previous one. It is adapted to the stratified spray situation. Namely, we have to produce sprays not on a single fiber but in a neighborhood of the fiber in each stratum (see Definition 5.4). In our case it will be on the whole stratum. The following definitions are straightforward. The first one was introduced in [Andrist and Kutzschebauch 2018, page 918].

Definition 9.2. Let $\pi : X \to Y$ be a holomorphic map between complex manifolds and denote by $d\pi : TX \to TY$ the tangent map. We call a holomorphic vector field θ on X *fiber-preserving* if $d\pi(\theta) = 0$.

Definition 9.3. A subset N of a complex manifold M is called invariant with respect to a collection of vector fields on M if for each of the vector fields we have: for each starting point $x \in N$ the local flow of the field (which is defined in a neighborhood of time 0) remains contained in N.

Lemma 9.4. *Consider a submersion $\pi : M \to Y$ with connected fibers $M_y := \pi^{-1}(y)$ and a finite collection of complete fiber-preserving holomorphic vector fields on M such that in each fiber M_y there is a point $x \in M_y$ where they span the tangent space $T_x M_y$. Suppose there is no analytic subset N of M contained in a fiber M_y which is invariant under the flows of $\theta_1, \ldots, \theta_k$. Then a finite subset of the set $\Gamma(\theta_1, \ldots, \theta_k)$ is spanning $T_x M_{\pi(x)}$ for all $x \in M$.*

Proof. Let $N \subset M$ be the set of points x where $\operatorname{span}\{(\theta_1, \ldots, \theta_k)\} \neq T_x M_{\pi(x)}$. By assumption $N \cup M_y$ is a proper analytic subset of M_y for each $y \in Y$. Since there is no invariant analytic subset different from the fibers for each $x_0 \in N$, there is a field θ_i whose flow starting in x_0 will leave N, i.e., go through points where $(\theta_1, \ldots, \theta_k)$ span $T_x M_{\pi(x)}$. Now choose (at most countably many) points, one from each component of N. As in the proof of the proceeding lemma find finitely many times t_i and enlarge the collection $\theta_1, \ldots, \theta_k$ by the pullbacks $(\alpha_i(t_i))^*(\theta_m)i,\ m = 1, \ldots, k$. We then get a new finite collection of complete fields where the set of points where this new collection does not span the tangent space of the π-fiber has smaller dimension. By finite induction on the dimension we get the desired result. □

10. Auxiliary quantities and results

Define

$$\mathcal{M}^K_{x_1x_2x_3} = \begin{pmatrix} \partial P^K_1/\partial x_1 & \partial P^K_1/\partial x_2 & \partial P^K_1/\partial x_3 \\ \partial P^K_2/\partial x_1 & \partial P^K_2/\partial x_2 & \partial P^K_2/\partial x_3 \\ \partial P^K_3/\partial x_1 & \partial P^K_3/\partial x_2 & \partial P^K_3/\partial x_3 \\ \partial P^K_4/\partial x_1 & \partial P^K_4/\partial x_2 & \partial P^K_4/\partial x_3 \end{pmatrix} \tag{10.0.1}$$

for any triple x_1, x_2, x_3 from $\vec{Z}_K$. Removing the j-th row from $\mathcal{M}^K_{x_1x_2x_3}$ gives us 3×3 matrices which we denote by $\mathcal{M}^{K,j}_{x_1x_2x_3}$. Let

$$\mathcal{R}^{K,j}_{x_1x_2x_3} = \det \mathcal{M}^{K,j}_{x_1x_2x_3}.$$

The significance of the functions $\mathcal{R}^{K,j}_{x_1x_2x_3}$ is understood if one notices, because of (8.0.1), that

$$\mathcal{R}^{2k+1,1}_{x_1x_2x_3} = \partial^{2k}_{x_1x_2x_3} P^{2k}_2, \tag{10.0.2}$$

$$\mathcal{R}^{2k+1,2}_{x_1x_2x_3} = \partial^{2k}_{x_1x_2x_3} P^{2k}_1 \tag{10.0.3}$$

and that

$$\mathcal{R}^{2k,3}_{x_1x_2x_3} = \partial^{2k-1}_{x_1x_2x_3} P_4^{2k-1}, \tag{10.0.4}$$

$$\mathcal{R}^{2k,4}_{x_1x_2x_3} = \partial^{2k-1}_{x_1x_2x_3} P_3^{2k-1}. \tag{10.0.5}$$

From (7.0.1) and (7.0.2) we get the relations

$$\begin{pmatrix} \mathcal{R}^{2k+1,1}_{x_1x_2x_3} \\ \mathcal{R}^{2k+1,2}_{x_1x_2x_3} \\ \mathcal{R}^{2k+1,3}_{x_1x_2x_3} \\ \mathcal{R}^{2k+1,4}_{x_1x_2x_3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -z_{3k+1} & z_{3k+2} & 1 & 0 \\ z_{3k+2} & -z_{3k+3} & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathcal{R}^{2k,1}_{x_1x_2x_3} \\ \mathcal{R}^{2k,2}_{x_1x_2x_3} \\ \mathcal{R}^{2k,3}_{x_1x_2x_3} \\ \mathcal{R}^{2k,4}_{x_1x_2x_3} \end{pmatrix} \tag{10.0.6}$$

and

$$\begin{pmatrix} \mathcal{R}^{2k,1}_{x_1x_2x_3} \\ \mathcal{R}^{2k,2}_{x_1x_2x_3} \\ \mathcal{R}^{2k,3}_{x_1x_2x_3} \\ \mathcal{R}^{2k,4}_{x_1x_2x_3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & -w_{3k-2} & w_{3k-1} \\ 0 & 1 & w_{3k-1} & -w_{3k} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathcal{R}^{2k-1,1}_{x_1x_2x_3} \\ \mathcal{R}^{2k-1,2}_{x_1x_2x_3} \\ \mathcal{R}^{2k-1,3}_{x_1x_2x_3} \\ \mathcal{R}^{2k-1,4}_{x_1x_2x_3} \end{pmatrix}. \tag{10.0.7}$$

Consider the vector fields $\theta^L_{x_1x_2x_3}$ and $\phi^L_{x_1x_2x_3}$, where $(x_1, x_2, x_3) \in \mathcal{T}_L$ and $3 \le L \le K$. Rewriting (8.0.12), (8.0.14), (8.0.15) and (8.0.16) using these functions we get

$$\begin{aligned} \theta^{2k+1}_{x_1x_2x_3} = P_3^{2k}(\vec{Z}_{2k})^2 \partial^{2k}_{x_1x_2x_3} + P_3^{2k}(\vec{Z}_{2k}) \mathcal{R}^{2k+1,1}_{x_1x_2x_3}(\vec{Z}_{2k}) \frac{\partial}{\partial z_{3k+2}} \\ + \big(P_3^{2k}(\vec{Z}_{2k}) \mathcal{R}^{2k+1,2}_{x_1x_2x_3}(\vec{Z}_{2k}) - P_4^{2k}(\vec{Z}_{2k}) \mathcal{R}^{2k+1,1}_{x_1x_2x_3}(\vec{Z}_{2k})\big) \frac{\partial}{\partial z_{3k+1}}, \end{aligned} \tag{10.0.8}$$

$$\begin{aligned} \phi^{2k+1}_{x_1x_2x_3} = P_4^{2k}(\vec{Z}_{2k})^2 \partial^{2k}_{x_1x_2x_3} + P_4^{2k}(\vec{Z}_{2k}) \mathcal{R}^{2k+1,2}_{x_1x_2x_3}(\vec{Z}_{2k}) \frac{\partial}{\partial z_{3k+2}} \\ + \big(P_4^{2k}(\vec{Z}_{2k}) \mathcal{R}^{2k+1,1}_{x_1x_2x_3}(\vec{Z}_{2k}) - P_3^{2k}(\vec{Z}_{2k}) \mathcal{R}^{2k+1,2}_{x_1x_2x_3}(\vec{Z}_{2k})\big) \frac{\partial}{\partial z_{3k+3}}, \end{aligned} \tag{10.0.9}$$

$$\begin{aligned} \theta^{2k}_{x_1x_2x_3} = P_1^{2k-1}(\vec{Z}_{2k-1})^2 \partial^{2k-1}_{x_1x_2x_3} + P_1^{2k-1}(\vec{Z}_{2k-1}) \mathcal{R}^{2k,3}_{x_1x_2x_3}(\vec{Z}_{2k-1}) \frac{\partial}{\partial w_{3k-1}} \\ + \big(P_1^{2k-1}(\vec{Z}_{2k-1}) \mathcal{R}^{2k,4}_{x_1x_2x_3}(\vec{Z}_{2k-1}) - P_2^{2k-1}(\vec{Z}_{2k-1}) \mathcal{R}^{2k,3}_{x_1x_2x_3}(\vec{Z}_{2k-1})\big) \frac{\partial}{\partial w_{3k-2}}, \end{aligned} \tag{10.0.10}$$

$$\begin{aligned} \phi^{2k}_{x_1x_2x_3} = P_2^{2k-1}(\vec{Z}_{2k-1})^2 \partial^{2k-1}_{x_1x_2x_3} + P_2^{2k-1}(\vec{Z}_{2k-1}) \mathcal{R}^{2k,4}_{x_1x_2x_3}(\vec{Z}_{2k-1}) \frac{\partial}{\partial w_{3k-1}} \\ + \big(P_2^{2k-1}(\vec{Z}_{2k-1}) \mathcal{R}^{2k,3}_{x_1x_2x_3}(\vec{Z}_{2k-1}) - P_1^{2k-1}(\vec{Z}_{2k-1}) \mathcal{R}^{2k,4}_{x_1x_2x_3}(\vec{Z}_{2k-1})\big) \frac{\partial}{\partial w_{3k}}. \end{aligned} \tag{10.0.11}$$

We see that half of the functions $\mathcal{R}^{K,j}_{x_1x_2x_3}$ occur in the coefficients of the last three directions. As already observed the fields $\theta^L_{x_1x_2x_3}$ and $\phi^L_{x_1x_2x_3}$ for $L < K$ have zero components along the last three directions. We have to make sure that the projection onto the last three variables of the collection of fields $\theta^K_{x_1x_2x_3}$ and $\phi^K_{x_1x_2x_3}$ spans a 3-dimensional space. If this is true for a point, we will say that the fields span all new directions in the point. In order to determine if our fields span all new directions in a point $\vec{Z}_K \in \mathcal{F}^K_{a_1a_2a_3a_4}$

we will use the following. Let $N_K = |\mathcal{T}_K|$ be the number of complete triples. Define the $(2 \times N_{K-1})$-matrices

$$\Omega^K_{x_1x_2x_3}(\vec{Z}_K) = \begin{cases} \begin{pmatrix} \mathcal{R}^{K-1,1}_{x_1x_2x_3}(\vec{Z}_K) & \cdots \\ \mathcal{R}^{K-1,2}_{x_1x_2x_3}(\vec{Z}_K) & \cdots \end{pmatrix} & \text{when } K \text{ odd,} \\ \begin{pmatrix} \mathcal{R}^{K-1,3}_{x_1x_2x_3}(\vec{Z}_K) & \cdots \\ \mathcal{R}^{K-1,4}_{x_1x_2x_3}(\vec{Z}_K) & \cdots \end{pmatrix} & \text{when } K \text{ even,} \end{cases}$$

where (x_1, x_2, x_3) run over all triples in $\mathcal{T}_{K-1}$. Using the formulas (10.0.8), (10.0.9), (10.0.10), (10.0.11), and remembering that a fiber $\mathcal{F}^K_{(a_1,a_2,a_3,a_4)}$ is called generic if $(a_1, a_2) \neq (0, 0)$ when K is even and if $(a_3, a_4) \neq (0, 0)$ when K is odd, it is an exercise in linear algebra to prove the lemma below.

Lemma 10.1. *If in a point $\vec{Z}_K \in \mathcal{F}^K_{a_1a_2a_3a_4}$ in a generic fiber*

$$\operatorname{Rank} \Omega^K_{x_1x_2x_3}(\vec{Z}_K) = 2$$

then

$$\{\theta^K_{x_1x_2x_3} : (x_1, x_2, x_3) \in \mathcal{T}_{K-1}\} \cup \{\phi^K_{x_1x_2x_3} : (x_1, x_2, x_3) \in \mathcal{T}_{K-1}\} \cup \{\gamma^K\}$$

span all three new directions. If

$$\operatorname{Rank} \Omega^K_{x_1x_2x_3}(\vec{Z}_K) = 1$$

then

$$\{\theta^K_{x_1x_2x_3} : (x_1, x_2, x_3) \in \mathcal{T}_{K-1}\} \cup \{\phi^K_{x_1x_2x_3} : (x_1, x_2, x_3) \in \mathcal{T}_{K-1}\} \cup \{\gamma^K\}$$

span two out of three new directions.

Because of the formulas (10.0.6) and (10.0.7) we have the lemma below.

Lemma 10.2. *Let $K \leq L$ and put*

$$\mathcal{M}^L_K(\vec{Z}_L) = \begin{pmatrix} \mathcal{R}^{L,1}_{x_1x_2x_3}(\vec{Z}_L) & \ldots \\ \mathcal{R}^{L,2}_{x_1x_2x_3}(\vec{Z}_L) & \ldots \\ \mathcal{R}^{L,3}_{x_1x_2x_3}(\vec{Z}_L) & \ldots \\ \mathcal{R}^{L,4}_{x_1x_2x_3}(\vec{Z}_L) & \ldots \end{pmatrix},$$

where (x_1, x_2, x_3) run over all triples in $\mathcal{T}_K$. For all $L \geq K$

$$\operatorname{Rank} \mathcal{M}^K_K(\vec{Z}_L) = \operatorname{Rank} \mathcal{M}^L_K(\vec{Z}_L).$$

The importance of Lemmas 10.1 and 10.2 is seen in the following corollary.

Corollary 10.3. *Let $L > K$ and $\vec{Z}_K$ be a point where* $\operatorname{Rank} \mathcal{M}^K_K(\vec{Z}_K) = 4$. *Then for all points $\vec{Z}_L$ contained in a generic fiber $\mathcal{F}^L_{(a_1,a_2,a_3,a_4)}$ such that $\pi(\vec{Z}_L) = \vec{Z}_K$, the complete fields*

$$\{\theta^L_{x_1x_2x_3} : (x_1, x_2, x_3) \in \mathcal{T}_{L-1}\} \cup \{\phi^L_{x_1x_2x_3} : (x_1, x_2, x_3) \in \mathcal{T}_{L-1}\} \cup \{\gamma^L\}$$

span all new directions (the directions along the last three variables in $(\mathbb{C}^3)^L$).

Proof. Two rows of the rank-4 matrix $\mathcal{M}^L_K(\vec{Z}_L)$ are linearly independent. □

$\partial^2_{x_1x_2x_3}$	$\mathcal{R}^{2,1}_{x_1x_2x_3}$	$\mathcal{R}^{2,2}_{x_1x_2x_3}$	$\mathcal{R}^{2,3}_{x_1x_2x_3}$	$\mathcal{R}^{2,4}_{x_1x_2x_3}$
$\partial^2_{w_1w_2w_3}$	0	0	0	0
$\partial^2_{z_2w_2w_3}$	0	z_3^2	0	0
$\partial^2_{z_3w_1w_2}$	z_2^2	0	0	0
$\partial^2_{z_2w_1w_3}$	0	z_2z_3	0	0
$\partial^2_{z_3w_1w_3}$	z_2z_3	0	0	0
$\partial^2_{z_2z_3w_1}$	z_2w_2	$-z_2w_3$	0	z_2
$\partial^2_{z_2z_3w_3}$	$-z_3w_1$	z_3w_2	z_3	0

Table 1. The expressions for $\mathcal{R}^{2,i}_{x_1x_2x_3}$.

Corollary 10.4. *Let $L \geq 3$ and $\vec{Z}_L$ be a point that is contained in a generic fiber $\mathcal{F}^L_{(a_1,a_2,a_3,a_4)}$ and such that $z_2z_3 \neq 0$. Then*

$$\{\theta^L_{x_1x_2x_3} : (x_1, x_2, x_3) \in \mathcal{T}_{L-1}\} \cup \{\phi^L_{x_1x_2x_3} : (x_1, x_2, x_3) \in \mathcal{T}_{L-1}\} \cup \{\gamma^L\}$$

span all new directions in $\vec{Z}_L$.

Proof. The corresponding matrix for $L = 3$ is contained in Table 1. From this table the claim is an easy exercise in linear algebra. □

In order to use this corollary we need the following lemma.

Lemma 10.5. *We have the following cases for the function $P = z_2z_3$ and the fibers $\mathcal{F}^K$:*

(1) *P is not identically zero on $\mathcal{F}^K_{(a_1,a_2,a_3,a_4)}$ for $K \geq 5$. For these K the fibers $\mathcal{F}^K_{(a_1,a_2,a_3,a_4)}$ are irreducible.*

(2) *The fibers $\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}$ are irreducible except when*

$$(a_1, a_2, a_3, a_4) = (0, 0, 0, 1).$$

The function P is not identically zero on fibers except for one component of $\mathcal{F}^4_{(0,0,0,1)}$.

(3) *The fibers $\mathcal{F}^3_{(a_1,a_2,a_3,a_4)}$ are irreducible except when*

$$(a_1, a_2, a_3, a_4) = (a_1, a_2, 0, 1).$$

The function P is not identically zero on fibers except on $\mathcal{F}^3_{(0,a_2,0,0)}$ or $\mathcal{F}^3_{(a_1,0,0,0)}$ or on one component of the reducible fiber $\mathcal{F}^3_{(a_1,a_2,0,1)}$ where it is identically zero.

Proof. We first prove (3). The fibers $\mathcal{F}^3_{(a_1,a_2,0,0)}$ are just biholomorphic to $\mathbb{C}^5$ and z_2, z_3 are constantly equal to a_1, a_2. This shows that they are irreducible and that the assertion about P is true. The fibers $\mathcal{F}^3_{(a_1,a_2,0,1)}$ are isomorphic to the variety $\mathcal{G}^2_{(0,1)}$ given by two equations which can be written in matrix form as

$$\begin{pmatrix} w_1 & w_2 \\ w_2 & w_3 \end{pmatrix} \begin{pmatrix} z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{10.0.12}$$

From this it can be seen that $\mathcal{G}^2_{(0,1)}$ has two irreducible components. One is

$$A_1 = \{z_2, z_3 : z_2{=}z_3{=}0\} \cong \mathbb{C}^3_{w_1w_2w_3} \tag{10.0.13}$$

and the other is

$$A_2 = \left\{ \begin{pmatrix} w_1 & w_2 \\ w_2 & w_3 \end{pmatrix} : \begin{pmatrix} w_1 & w_2 \\ w_2 & w_3 \end{pmatrix} \begin{pmatrix} z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ and } \det \begin{pmatrix} w_1 & w_2 \\ w_2 & w_3 \end{pmatrix} = 0 \right\}. \tag{10.0.14}$$

The singularity set of $\mathcal{G}^2_{(0,1)}$ is $A_1 \cap A_2$. Clearly P is identically zero on A_1 and not identically zero on A_2. Observe that $\mathcal{F}^3_{(a_1,a_2,0,1)}$ are connected, their smooth part consists of the two connected components $A_1 \setminus A_2$ and $A_2 \setminus A_1$.

The smooth generic fibers $\mathcal{F}^3_{(a_1,a_2,a_3,a_4)}$ for $(a_3, a_4) \notin \{(0, 0), (0, 1)\}$ are isomorphic to the variety $\mathcal{G}^2_{(a_3,a_4)}$ given by the two equations

$$z_2 w_1 + z_3 w_2 = a_3, \tag{10.0.15}$$

$$z_2 w_2 + z_3 w_3 + 1 = a_4. \tag{10.0.16}$$

In the case $z_2 \neq 0$ these equations can be used to express w_2 and w_3 by the other variables and we get a chart isomorphic to $\mathbb{C}^\star_{z_2} \times \mathbb{C}_{z_3} \times \mathbb{C}_{w_3}$. In the case $z_3 \neq 0$ we can express w_2 and w_3 by other variables, which gives us a similar chart. Thus $\mathcal{G}^2_{(a_3,a_4)}$ is covered by two connected charts with nonempty intersection which shows that it is connected. Thus the smooth generic fibers $\mathcal{F}^3_{(a_1,a_2,a_3,a_4)}$ are irreducible. The function P is not identically zero on both charts. The assertion (3) is completely proven.

Next we prove assertion (2). The nongeneric fibers $\mathcal{F}^4_{(0,0,a_3,a_4)}$ are isomorphic to $\mathcal{F}^3_{(0,0,a_3,a_4)} \times \mathbb{C}^3$, where $\mathbb{C}^3$ corresponds to the new variables w_4, w_5, w_6. All assumptions about these fibers follow therefore from the corresponding assumptions about $\mathcal{F}^3_{(0,0,a_3,a_4)}$.

In the case of generic fibers which are known to be smooth (see Section 7) we just have to prove that they are connected. For this consider

$$\mathcal{F}^4_{(a_1,a_2,a_3,a_4)} = \bigcup_{(w_4,w_5,w_6)\in\mathbb{C}^3} \mathcal{F}^3_{(a_1,a_2,b_3,b_4)}, \tag{10.0.17}$$

where $b_3 = a_3 - w_4 a_1 - w_5 a_2$ and $b_4 = a_4 - w_5 a_1 - w_6 a_2$. In other words we consider the surjective projection $\rho : \mathcal{F}^4_{(a_1,a_2,a_3,a_4)} \to \mathbb{C}^3$, mapping a point to its last three coordinates (w_4, w_5, w_6), where the ρ-fibers are just fibers $\mathcal{F}^3_{(a_1,a_2,b_3,b_4)}$. Connectedness of the ρ-fibers implies that a connected component of $\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}$ has to be ρ-saturated. Since ρ is a submersion in generic points of the fiber (it is not a submersion only in singular points of an $\mathcal{F}^3$-fiber), any connected component of $\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}$ is equal to $\rho^{-1}(U)$, where U is some open subset of the base $\mathbb{C}^3$. Since the base is connected and ρ is surjective, connectedness of $\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}$ follows. The function P is not identically zero on any $\mathcal{F}^3$-fiber contained in $\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}$, and thus not identically zero on $\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}$ itself. This concludes the proof of (2).

Last we prove assertion (1). The connectedness of the fibers $\mathcal{F}^K_{(a_1,a_2,a_3,a_4)}$ for $K \geq 5$ can be proven by induction in a way similar to the connectedness of the generic $\mathcal{F}^4$-fibers is deduced from the properties of $\mathcal{F}^3$-fibers. As above we consider the surjective projection $\rho : \mathcal{F}^K_{(a_1,a_2,a_3,a_4)} \to \mathbb{C}^3$ onto the last three variables whose fibers are $\mathcal{F}^{K-1}$-fibers. Since again $\mathcal{F}^{K-1}$-fibers are connected and ρ is a submersion in smooth points of the $\mathcal{F}^{K-1}$-fibers, any connected component of $\mathcal{F}^K_{(a_1,a_2,a_3,a_4)}$ is of the form $\rho^{-1}(U)$, where U is some open subset of the base $\mathbb{C}^3$.

In addition we will prove by induction that the smooth part of the singular fibers

$$\mathcal{F}^K_{(a_1,a_2,a_3,a_4)} \setminus \operatorname{Sing} \mathcal{F}^K_{(a_1,a_2,a_3,a_4)}$$

is connected for $K \geq 5$. Together with connectedness of the fibers this implies the irreducibility of the fibers.

For even K, the singular fibers are the singular $\mathcal{F}^{K-1}$-fibers times $\mathbb{C}^3$ and therefore the connectedness of the smooth part follows by the induction hypothesis.

For odd $K = 2k+1$, we are faced with the following situation: The singular fiber is $\mathcal{F}^K_{(a_1,a_2,0,1)}$ and it is fibered by $\mathcal{F}^{K-1}$-fibers all of which are smooth except for the fibers $\mathcal{F}^{K-1}_{(0,0,0,1)}$. The union of those fibers forms a codimension-2 subvariety of $\mathcal{F}^K_{(a_1,a_2,0,1)}$ (given by the equations $z_{3k+2} = z_{3k+3} = 0$). By the argument above, the complement, call it W, of this union in $\mathcal{F}^K_{(a_1,a_2,0,1)}$ is connected. The singular points of $\mathcal{F}^K_{(a_1,a_2,0,1)}$ are contained in that union and are contained in (but not equal to) the union of the singular points of the fibers $\mathcal{F}^{K-1}_{(0,0,0,1)}$. We want to prove that any smooth point p of $\mathcal{F}^K_{(a_1,a_2,0,1)}$ which is contained in a fiber $\mathcal{F}^{K-1}_{(0,0,0,1)}$ is contained in the connected component containing W. Since the complement of W has codimension 2 in $\mathcal{F}^K_{(a_1,a_2,0,1)}$, an open neighborhood of p in $\mathcal{F}^K_{(a_1,a_2,0,1)}$ has to intersect W, which gives the desired conclusion.

As in the proof of (2), the function P cannot be identically zero on any fiber $\mathcal{F}^K_{(a_1,a_2,a_3,a_4)}$ since this fiber contains $\mathcal{F}^{K-1}$-fibers on which, by the induction hypothesis, P is not identically zero. $\square$

Remark 10.6. The fact that after a certain number of factors the fibers of the fibration all become irreducible is very general. It was proven by J. Draisma as an outcome of an interesting discussion with the second author. The irreducibility statement in our lemma is just an example of a much more general property. We refer the interested reader to [Draisma 2022]. The exact number at and past which irreducibility of the fibers holds (in our case 5) is not known in general, although Draisma gives a bound.

Definition 10.7. Let M be a manifold and A be a set of complete vector fields on M. The flows of elements of A give one-parameter subgroups of $\mathrm{Aut}(M)$. Denote by S the group generated by elements of those one-parameter subgroups (finite compositions of time maps of vector fields of elements from A). Define

$$\Gamma(A) = \{\alpha^* X : \alpha \in S \text{ and } X \in A\}.$$

Obviously $\Gamma(A)$ consists of complete vector fields and we call it *the collection generated by A*.

Definition 10.8. Let $L \geq 3$. We define

$$\mathcal{Q}_L = \Gamma\Bigg(\bigcup_{J=3}^{L} \big\{\{\theta^J_{x_1x_2x_3} : (x_1,x_2,x_3) \in \Xi_J\} \cup \{\phi^J_{x_1x_2x_3} : (x_1,x_2,x_3) \in \Xi_J\} \cup \{\gamma^J\}\big\}\Bigg).$$

At each step of the induction we will prove the following proposition, which plays a crucial role in the inductive proof of Proposition 3.6.

Proposition 10.9. *For each $L \geq 4$ we have: There are finitely many (complete) fields from $\mathcal{Q}_L$ which span the tangent space $T_x\mathcal{F}^L$ at each smooth point of any generic fiber $\mathcal{F}^L$. For $L = 3$ there are finitely many (complete) fields from $\mathcal{Q}_3$ which span the tangent space $T_x\mathcal{F}^3$ at each point of any smooth generic fiber $\mathcal{F}^3$.*

Remark 10.10. For $L = 3$ singular generic fibers $\mathcal{F}^3_{(a_1,a_2,0,1)}$ have two irreducible components and we can prove the statement about smooth points on generic fibers only for one of those components. It is false for the other component.

11. Proof of Proposition 3.6: three matrix factors

In Table 2 we list the coefficients of the fields $\partial^2_{x_1x_2x_3}$ for all $x_1, x_2, x_3 \in \mathcal{T}_2$.

We first consider the stratum of *smooth generic fibers*, where we have

$$(a_3, a_4) \notin \{(0, 0), (0, 1)\}.$$

Notice that $z_2 = z_3 = 0$ is contained in $\mathcal{F}^3_{(z_5,z_6,0,1)}$ and therefore z_2 and z_3 are never simultaneously zero on any fiber in this stratum. It is enough to show that $\mathcal{G}^2_{(a_3,a_4)}$ is elliptic. We see from the table that the fields $\partial^2_{z_3w_1w_3}$, $\partial^2_{z_2w_1w_3}$, $\partial^2_{w_1w_2w_3}$ span the tangent space $T_{\vec{Z}_2}\mathcal{G}^2_{(a_3,a_4)}$ for all points $\vec{Z}_2$ where $z_2z_3 \neq 0$. The complement of this good set is the disjoint union of the analytic subsets $\mathcal{A} = \{\vec{Z}_2 : z_2 = 0\}$ and $\mathcal{B} = \{\vec{Z}_2 : z_3 = 0\}$. From the table we see that $\partial^2_{z_2w_2w_3}(z_2) = z_3^2$, which is nowhere-zero on $\mathcal{A}$. Also $\partial^2_{z_3w_1w_2}(z_3) = z_2^2$, which is nowhere-zero on $\mathcal{B}$. By Lemma 9.1 there exist finitely many complete fields from

$$\Gamma(\{\partial^2_{x_1x_2x_3} : (x_1, x_2, x_3) \in \mathcal{T}_2\})$$

that span the tangent space $T_{\vec{Z}_2}\mathcal{G}^2_{(a_3,a_4)}$ for all points in the stratum. Therefore $\mathcal{G}^2_{(a_3,a_4)}$ is elliptic. It follows that there are finitely many complete fields from

$$\Gamma\big(\{\theta^3_{x_1x_2x_3} : (x_1, x_2, x_3) \in \mathcal{T}_2\} \cup \{\phi^3_{x_1x_2x_3} : (x_1, x_2, x_3) \in \mathcal{T}_2\} \cup \{\gamma^3\}\big)$$

that span the tangent space $T_{\vec{Z}_3}\mathcal{F}^3_{(a_1,a_2,a_3,a_4)}$ for all points in the stratum.

Now we consider the stratum of *nonsmooth generic fibers* $(a_3, a_4) = (0, 1)$. The two equations defining $\mathcal{G}^2_{(0,1)}$ can be written in matrix form as

$$\begin{pmatrix} w_1 & w_2 \\ w_2 & w_3 \end{pmatrix} \begin{pmatrix} z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{11.0.1}$$

	$\partial/\partial z_2$	$\partial/\partial z_3$	$\partial/\partial w_1$	$\partial/\partial w_2$	$\partial/\partial w_3$
$\partial^2_{w_1w_2w_3}$	0	0	z_3^2	$-z_2z_3$	z_2^2
$\partial^2_{z_2w_2w_3}$	z_3^2	0	0	$-w_1z_3$	$w_1z_2-w_2z_3$
$\partial^2_{z_3w_1w_2}$	0	z_2^2	$z_3w_3-w_2z_2$	$-z_2w_3$	0
$\partial^2_{z_2w_1w_3}$	z_2z_3	0	$-w_1z_3$	0	$-z_2w_2$
$\partial^2_{z_3w_1w_3}$	0	z_2z_3	$-w_2z_3$	0	$-z_2w_3$
$\partial^2_{z_2z_3w_1}$	$-z_2w_3$	z_2w_2	$w_1w_3-w_2^2$	0	0
$\partial^2_{z_2z_3w_3}$	w_2z_3	w_1z_3	0	0	$w_1w_3-w_2^2$

Table 2. Coefficients of complete vector fields. For example, $\partial^2_{w_1w_2w_3} = z_3^2(\partial/\partial w_1) - z_2z_3(\partial/\partial w_2) + z_2^2(\partial/\partial w_3)$.

Recall that $\mathcal{G}^2_{(0,1)}$ has two irreducible components. The components are given by (see (10.0.13) and (10.0.14))

$$A_1 = \{z_2, z_3 : z_2 = z_3 = 0\} \cong \mathbb{C}^3_{w_1w_2w_3} \tag{11.0.2}$$

and

$$A_2 = \left\{ \begin{pmatrix} w_1 & w_2 \\ w_2 & w_3 \end{pmatrix} : \begin{pmatrix} w_1 & w_2 \\ w_2 & w_3 \end{pmatrix} \begin{pmatrix} z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ and } \det \begin{pmatrix} w_1 & w_2 \\ w_2 & w_3 \end{pmatrix} = 0 \right\}. \tag{11.0.3}$$

The singularity set of $\mathcal{G}^2_{(0,1)}$ is $A_1 \cap A_2$. We have to show that the smooth part of $\mathcal{G}^2_{(0,1)}$, that is, the disjoint union of $A_1 \setminus A_2$ and $A_2 \setminus A_1$, is elliptic. In the proof for the smooth generic case it is shown that on the set where z_2 and z_3 are not both zero, there exists a collection of complete spanning vector fields. Since $A_2 \setminus A_1$ is contained in that set, we need only consider $A_1 \setminus A_2$. The set $A_1 \setminus A_2$ is biholomorphic to $\mathbb{C}^3 \setminus \{w_1w_3 - w_2^2 = 0\}$. The vector fields $(w_1w_3 - w_2^2)(\partial/\partial w_1)$, $(w_1w_3 - w_2^2)(\partial/\partial w_3)$, $2w_2(\partial/\partial w_1) + w_3(\partial/\partial w_2)$, $2w_2(\partial/\partial w_3) + w_1(\partial/\partial w_2)$ are complete on $\mathbb{C}^3 \setminus \{w_1w_3 - w_2^2 = 0\}$ and span the tangent space in all points outside the analytic set $A' = \{w_1 = w_3 = 0\} \cap (A_1 \setminus A_2)$. Since w_2 is nowhere-zero on A' any of the four complete fields points out of A'. By Lemma 9.1 the proof is complete. Observe that we also have proved Proposition 10.9 for $L = 3$. Notice that the fields $2w_2(\partial/\partial w_1) + w_3(\partial/\partial w_2)$, $2w_2(\partial/\partial w_3) + w_1(\partial/\partial w_2)$ are not in $\mathcal{Q}_3$ and this explains the difference between $L = 3$ and $L \geq 4$ in Proposition 10.9. See Remark 10.10.

The stratum of *nongeneric fibers* is a locally trivial bundle with fibers $\mathbb{C}^5 (\cong \mathcal{F}^3_{(a_1,a_2,0,0)})$ which is an elliptic submersion.

12. Proof of Proposition 3.6: four matrix factors

We begin the proof by studying the stratum of *generic fibers*, $(a_1, a_2) \neq (0, 0)$. We write

$$\mathcal{F}^4_{(a_1,a_2,a_3,a_4)} = \bigcup_{(w_4,w_5,w_6) \in \mathbb{C}^3} \mathcal{F}^3_{(a_1,a_2,b_3,b_4)}, \tag{12.0.1}$$

where $b_3 = a_3 - w_4a_1 - w_5a_2$ and $b_4 = a_4 - w_5a_1 - w_6a_2$. We need to find finitely many complete vector fields spanning $T_{\vec{Z}_4}\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}$ for points $\vec{Z}_4$ in the stratum of generic fibers. Because of (12.0.1) there are b_3 and b_4 so that $\vec{Z}_3 \in \mathcal{F}^3_{(a_1,a_2,b_3,b_4)}$ and $\vec{Z}_4 = (\vec{Z}_3, w_4, w_5, w_6)$. We first consider the set of points in these fibers having the property that $(b_3, b_4) \neq (0, 0)$ or $(0, 1)$. Under these assumptions, $\vec{Z}_4$ lies in a generic smooth fiber $\mathcal{F}^3_{(a_1,a_2,b_3,b_4)}$ and we know from Section 11 that there is a finite collection of fields from $\mathcal{Q}_3$ which spans

$$T_{\vec{Z}_4}\mathcal{F}^3_{(a_1,a_2,b_3,b_4)} \subset T_{\vec{Z}_4}\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}.$$

Corollary 10.4 together with Lemma 10.5(3) shows that for the set defined by $z_2z_3 \neq 0$ (which is a Zariski open and dense set of points of the generic fiber $\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}$) the fields

$$\{\theta^4_{x_1x_2x_3} : (x_1, x_2, x_3) \in \mathcal{T}_3\} \cup \{\phi^4_{x_1x_2x_3} : (x_1, x_2, x_3) \in \mathcal{T}_3\} \cup \{\gamma^4\}$$

span the new directions w_4, w_5, w_6. Since these new directions are complementary to

$$T_{\vec{Z}_4}\mathcal{F}^3_{(a_1,a_2,b_3,b_4)} \subset T_{\vec{Z}_4}\mathcal{F}^4_{(a_1,a_2,a_3,a_4)},$$

we have found finitely many complete fields spanning $T_{\vec{Z}_4}\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}$ for points in a Zariski open dense set in all smooth generic fibers $\mathcal{F}^3_{(a_1,a_2,b_3,b_4)}$. Using Lemma 9.1 we get finitely many complete fields spanning $T_{\vec{Z}_4}\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}$ for all points in all generic fibers $\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}$ with the property that $(b_3,b_4)\neq(0,0)$ or $(0,1)$. Next we consider points $\vec{Z}_4$ where $(b_3,b_4)=(0,1)$, i.e,

$$\vec{Z}_4\in\mathcal{F}^3_{(a_1,a_2,0,1)}\subset\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}.$$

Remember that

$$\mathcal{F}^3_{(a_1,a_2,0,1)}=A_1\cup A_2=A_1\dot{\cup}(A_2\setminus A_1) \tag{12.0.2}$$

(see (10.0.13) and (10.0.14)), where A_1 and A_2 are irreducible components. In the proof for $K=3$ we saw that there is a finite collection from $\mathcal{Q}_3$ which spans all tangent spaces

$$T_{\vec{Z}_4}\mathcal{F}^3_{(a_1,a_2,0,1)}\subset T_{\vec{Z}_4}\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}$$

for all points in $A_2\setminus A_1$. Lemma 10.5(3) gives that z_2z_3 is not identically zero on $A_2\setminus A_1$ and as above, appealing to Lemma 9.1, we get spanning fields for the fiber $\mathcal{F}^4_{(a_1,a_2,a_3,a_4)}$ in all points of $A_2\setminus A_1$. Our aim is to exclude the existence of a subset of the fiber invariant under the flows of fields from $\mathcal{Q}_4$. By the reasoning above, such a subset must be contained in A_1 or the set of points $\vec{Z}_4$ where $(b_3,b_4)=(0,0)$. Next we show that such a subset is disjoint from A_1. A calculation shows that

$$\partial^3_{z_2z_3z_6}=(z_4w_2+z_5w_3)\frac{\partial}{\partial z_2}-(1+z_4w_1+z_5w_2)\frac{\partial}{\partial z_3}+\cdots.$$

Therefore the complete fields

$$\begin{aligned}\theta^4_{z_2z_3z_6}&=a_1^2\partial^3_{z_2z_3z_6}+\cdots,\\ \phi^4_{z_2z_3z_6}&=a_2^2\partial^3_{z_2z_3z_6}+\cdots\end{aligned}$$

move points out of A_1 (into the big orbit) unless, in addition to $z_2=z_3=0$, also

$$1+z_4w_1+z_5w_2=z_4w_2+z_5w_3=0. \tag{12.0.3}$$

Points in an invariant subset must also satisfy these equations. A calculation gives that $\partial^3_{z_4z_5z_6}=\partial/\partial z_4$ when $z_2=z_3=0$. Therefore the complete fields

$$\begin{aligned}\theta^4_{z_4z_5z_6}&=a_1^2\partial^3_{z_4z_5z_6}+\cdots,\\ \phi^4_{z_4z_5z_6}&=a_2^2\partial^3_{z_4z_5z_6}+\cdots\end{aligned}$$

move points out of this set since

$$\begin{aligned}\theta^4_{z_4z_5z_6}(1+z_4w_1+z_5w_2)&=a_1^2w_1,\\ \theta^4_{z_4z_5z_6}(z_4w_2+z_5w_3)&=a_1^2w_2,\\ \phi^4_{z_4z_5z_6}(1+z_4w_1+z_5w_2)&=a_2^2w_1,\\ \phi^4_{z_4z_5z_6}(z_4w_2+z_5w_3)&=a_2^2w_2\end{aligned}$$

cannot all be zero, because this would contradict (12.0.3). We now turn to points $\vec{Z}_4$ where $(b_3,b_4)=(0,0)$ and again show that these points are not contained in an invariant subset and hence no such invariant subset exists. We will find fields $\theta^4_{x_1x_2x_3}$ or $\phi^4_{x_1x_2x_3}$ such that $\mathcal{R}^{3,3}_{x_1x_2x_3} \neq 0$ or $\mathcal{R}^{3,4}_{x_1x_2x_3} \neq 0$. We begin by noticing that at points $\vec{Z}_4$ with $z_2z_3 \neq 0$ we can leave the invariant set. Also $z_2 = 0 = z_3$ cannot occur in $\mathcal{F}^3_{(a_1,a_2,0,0)}$. Two cases, $z_2 \neq 0 = z_3$ and $z_2 = 0 \neq z_3$, remain. Assume first that $z_2 \neq 0 = z_3$ (and $b_3 = b_4 = 0$). Here we begin by choosing the triple (z_3, w_1, w_2). Since $\mathcal{R}^{3,3}_{z_3w_1w_2} = -z_2^2z_4$ and $\mathcal{R}^{3,4}_{z_3w_1w_2} = z_2^2z_5$, we move out of $\mathcal{F}^3_{(a_1,a_2,0,0)}$ unless $z_4 = z_5 = 0$. Assuming in addition that $z_4 = z_5 = 0$ we choose the triple (z_2, z_3, w_1). For such points, $\mathcal{R}^{3,4}_{z_2z_3w_1} = z_2 + z_2w_3z_6$ (and $\mathcal{R}^{3,3}_{z_2z_3w_1} = 0$) so if $1 + w_3z_6 \neq 0$ we move out of $\mathcal{F}^3_{(a_1,a_2,0,0)}$. Choose (z_2, z_3, z_6). Notice that

$$\theta^4_{z_2z_3z_6} = a_1^2\frac{\partial}{\partial z_6} \quad \text{and} \quad \phi^4_{z_2z_3z_6} = a_2^2\frac{\partial}{\partial z_6}$$

at these points and $\theta^4_{z_2z_3z_6}(1 + w_3z_6) = a_1^2w_3$, $\phi^4_{z_2z_3z_6}(1 + w_3z_6) = a_2^2w_3$ which both cannot be zero since $1 + w_3z_6 = 0$ implies $w_3 \neq 0$ and we assume that $(a_1, a_2) \neq (0, 0)$.

Now assume that $z_2 = 0 \neq z_3$ (and also $b_3 = b_4 = 0$) and choose the triple (z_2, w_2, w_3). Since $\mathcal{R}^{3,3}_{z_2w_2w_3} = z_3^2z_5$ and $\mathcal{R}^{3,4}_{z_2w_2w_3} = -z_3^2z_6$ we move out of $\mathcal{F}^3_{(a_1,a_2,0,0)}$ unless $z_5 = z_6 = 0$. Assuming in addition that $z_5 = z_6 = 0$ we choose the triple (z_2, z_3, w_3). For such points $\mathcal{R}^{3,3}_{z_2z_3w_3} = z_3 + z_3w_1z_4$ (and $\mathcal{R}^{3,4}_{z_2z_3w_3} = 0$) so if $1 + w_1z_4 \neq 0$ we move out of $\mathcal{F}^3_{(a_1,a_2,0,0)}$. Therefore assume also that $1 + w_1z_4 = 0$ Choose (z_2, z_3, z_4). Notice that

$$\theta^4_{z_2z_3z_4} = a_1^2\frac{\partial}{\partial z_4} \quad \text{and} \quad \phi^4_{z_2z_3z_4} = a_2^2\frac{\partial}{\partial z_4}$$

at these points and $\theta^4_{z_2z_3z_4}(1 + w_1z_4) = a_1^2w_1$, $\phi^4_{z_2z_3z_4}(1 + w_1z_4) = a_2^2w_1$, which both cannot be zero since $1 + w_1z_4 = 0$ implies $w_1 \neq 0$ and we assumed that $(a_1, a_2) \neq (0, 0)$. This lets us conclude that there is no invariant subset with respect to $\mathcal{Q}_4$ and we have handled the stratum of generic fibers. Note that this proves Proposition 10.9 for $K = 4$.

We need to study the stratum of *nongeneric fibers*. This stratum consists of those fibers where $a_1 = a_2 = 0$. We notice that these fibers satisfy

$$\mathcal{F}^4_{(0,0,a_3,a_4)} = \mathcal{F}^3_{(0,0,a_3,a_4)} \times \mathbb{C}^3$$

and since $\mathcal{F}^3_{(0,0,a_3,a_4)}$ is elliptic we have proven Proposition 3.6 for $K = 4$.

13. Proof of Proposition 3.6: five matrix factors

We assume that $K = 5$ and we have seen that the submersions $\Phi_L = \pi_4 \circ \Psi_L$ are stratified elliptic submersions when $3 \leq L \leq 4$ and that Proposition 10.9 is true when $3 \leq L \leq 4$.

We study

$$\mathcal{F}^5_{(a_1,a_2,a_3,a_4)} = \bigcup_{(z_7,z_8,z_9)\in\mathbb{C}^3} \mathcal{F}^4_{(b_1,b_2,a_3,a_4)}, \tag{13.0.1}$$

where $b_1 = a_1 - z_7a_3 - z_8a_4$ and $b_2 = a_2 - z_8a_3 - z_9a_4$. Let $\vec{Z}_5 \in \mathcal{F}^5_{(a_1,a_2,a_3,a_4)}$. Because of (13.0.1) there are b_1 and b_2 so that $\vec{Z}_4 \in \mathcal{F}^4_{(b_1,b_2,a_3,a_4)}$ and $\vec{Z}_5 = (\vec{Z}_4, z_7, z_8, z_9)$.

First we study the stratum of *smooth generic fibers*. Fibers in this stratum are those satisfying $(a_3, a_4) \notin \{(0, 0), (0, 1)\}$. First notice that if $(b_1, b_2) \neq (0, 0)$ then $\mathcal{F}^4_{(b_1,b_2,a_3,a_4)}$ is a generic smooth fiber

	$\partial/\partial z_2$	$\partial/\partial z_3$	$\partial/\partial z_5$	$\partial/\partial z_6$
P_1^4	$1+w_1z_4$	w_3z_4	1	0
P_2^4	0	1	0	1
P_3^4	$w_1+w_4+w_1w_4z_4$	$w_2+w_5+w_2w_4z_4$	w_4	w_5
P_4^4	$w_2+w_5+w_1w_5z_4$	$w_3+w_6+w_2w_5z_4$	w_5	w_6

Table 3. The nonzero partial derivatives of P_1^4, P_2^4, P_3^4, and P_4^4.

for Φ_4 and as above, Proposition 10.9 for $L=4$, Corollary 10.4 and Lemma 10.5(2) show that for these points we have spanning fields. If $(b_1,b_2)=(0,0)$ then $\mathcal{F}^4_{(0,0,a_3,a_4)}$ is a nongeneric smooth fiber for Φ_4 and

$$\mathcal{F}^4_{(0,0,a_3,a_4)}=\mathcal{F}^3_{(0,0,a_3,a_4)}\times\mathbb{C}^3.$$

Since we assume that $(a_3,a_4)\neq(0,1)$ (in this case $(a_3,a_4)\neq(0,0)$ is automatic), we know by Corollary 10.4, Lemma 10.5(3) and Proposition 10.9 that we have spanning fields. This also shows that Proposition 10.9 holds for these fibers when $L=5$.

We now study the stratum of *singular generic fibers*. Here $(a_3,a_4)=(0,1)$. Again notice that if $(b_1,b_2)\neq(0,0)$ then

$$\mathcal{F}^4_{(b_1,b_2,0,1)}$$

is a generic smooth fiber for Φ_4, and Proposition 10.9 (for $L=4$), Corollary 10.4 and Lemma 10.5 show that for these points we have spanning fields as above. Next we study the case $(b_1,b_2)=(0,0)$. In this case we see that

$$\mathcal{F}^4_{(0,0,0,1)}\cong\mathcal{F}^3_{(0,0,0,1)}\times\mathbb{C}^3.$$

We write, as in Section 11,

$$\mathcal{F}^3_{(0,0,0,1)}=A_1\cup A_2.$$

In $A_2\setminus A_1$ we can use the argument as in the smooth generic case in Section 11: $z_2z_3\not\equiv 0$ and $\partial^2_{z_2w_2w_3}(z_2)=z_3^2$ make it possible to leave the set where $z_2=0$, and $\partial^2_{z_3w_1w_2}(z_3)=z_2^2$ makes it possible to leave the set where $z_3=0$.

Now we need to deal with points in $A_1\times\mathbb{C}^3\subset\mathcal{F}^4_{(0,0,0,1)}$. Because of the inclusion we find $z_5=z_6=0$. Define $C=\{z_2=z_3=z_5=z_6=0\}\subset\mathcal{F}^4_{(0,0,0,1)}\subset\mathcal{F}^5_{(0,0,0,1)}$, which contains the set of singularities

$$\operatorname{Sing}(\mathcal{F}^5_{(0,0,0,1)})=C\cap\left\{\operatorname{Rank}\begin{pmatrix}w_1&w_2&w_4&w_5\\w_2&w_3&w_5&w_6\end{pmatrix}<2\right\}.$$

In order to prove Propositions 10.9 and 3.6 we need to show that fields from $\mathcal{Q}_5$ move out from $C\setminus\operatorname{Sing}(\mathcal{F}^5_{(0,0,0,1)})$. Calculating the partial derivatives of $P_1^4,\dots,P_4^4$ in points of C we find that the ones that are nonzero are those listed in Table 3. We examine the complete field $\phi^5_{z_2z_3z_6}$. This field has some complicated components which on C take the form

$$\phi^5_{z_2z_3z_6}=\mathcal{D}_1\frac{\partial}{\partial z_2}+\mathcal{D}_2\frac{\partial}{\partial z_3}+\mathcal{D}_3\frac{\partial}{\partial z_6}+\cdots,$$

where

$$\mathcal{D}_1 = \det\begin{pmatrix} w_2+w_5+w_2w_4z_4 & w_5 \\ w_3+w_6+w_2w_5z_4 & w_6 \end{pmatrix},$$

$$\mathcal{D}_2 = \det\begin{pmatrix} w_1+w_4+w_1w_4z_4 & w_5 \\ w_2+w_5+w_1w_5z_4 & w_6 \end{pmatrix},$$

$$\mathcal{D}_3 = \det\begin{pmatrix} w_1+w_4+w_1w_4z_4 & w_2+w_5+w_2w_4z_4 \\ w_2+w_5+w_1w_5z_4 & w_3+w_6+w_2w_5z_4 \end{pmatrix}.$$

Whenever at least one of $\mathcal{D}_1$, $\mathcal{D}_2$ or $\mathcal{D}_3$ is nonzero we can move out of C. Suppose we are in a point of $C \setminus \mathrm{Sing}(\mathcal{F}^5_{(0,0,0,1)})$ where $\mathcal{D}_1 = \mathcal{D}_2 = \mathcal{D}_3 = 0$. Observe that

$$\mathrm{Rank}\begin{pmatrix} w_1 & w_2 & w_4 & w_5 \\ w_2 & w_3 & w_5 & w_6 \end{pmatrix} = \mathrm{Rank}\begin{pmatrix} w_1+w_4+w_1w_4z_4 & w_2+w_5+w_2w_4z_4 & w_4 & w_5 \\ w_2+w_5+w_1w_5z_4 & w_3+w_6+w_2w_5z_4 & w_5 & w_6 \end{pmatrix}$$

(in this case it is 2) since

$$\begin{pmatrix} w_1+w_4+w_1w_4z_4 \\ w_2+w_5+w_1w_5z_4 \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} + (1+w_1z_4)\begin{pmatrix} w_4 \\ w_5 \end{pmatrix},$$

$$\begin{pmatrix} w_2+w_5+w_2w_4z_4 \\ w_3+w_6+w_2w_5z_4 \end{pmatrix} = \begin{pmatrix} w_2 \\ w_3 \end{pmatrix} + \begin{pmatrix} w_5 \\ w_6 \end{pmatrix} + w_2z_4\begin{pmatrix} w_4 \\ w_5 \end{pmatrix}.$$

The fact that $\mathcal{D}_1 = \mathcal{D}_2 = \mathcal{D}_3 = 0$ means that the rank drops when we remove the third column from these matrices. This implies that the third column is nonzero and the other columns are multiples of a nonzero vector v which moreover is linearly independent of the third column. Now we use the field γ^3 (see (8.0.13)) to show that the set

$$I = C \setminus \mathrm{Sing}(\mathcal{F}^5_{(0,0,0,1)}) \cap \{\mathcal{D}_1 = \mathcal{D}_2 = \mathcal{D}_3 = 0\}$$

does not contain an invariant subset under fields from $\mathcal{Q}_5$. In the points of I we have that $\gamma^3 = \partial/\partial z_4$. We consider two cases.

<u>Case 1</u>: $(w_5, w_6) \neq (0,0)$. In this case

$$\det\begin{pmatrix} w_4 & w_5 \\ w_5 & w_6 \end{pmatrix} \neq 0.$$

We have

$$\gamma^3(\mathcal{D}_1) = w_2 \det\begin{pmatrix} w_4 & w_5 \\ w_5 & w_6 \end{pmatrix}.$$

Thus γ^3 moves points out of I unless $w_2 = 0$. Looking at

$$\gamma^3(\mathcal{D}_2) = w_1 \det\begin{pmatrix} w_4 & w_5 \\ w_5 & w_6 \end{pmatrix}$$

we see that $w_1 = 0$ for I to be invariant. Assuming in addition $w_1 = w_2 = 0$ we find that

$$\mathcal{D}_2 = \det\begin{pmatrix} w_4 & w_5 \\ w_5 & w_6, \end{pmatrix}$$

which is a contradiction since $\mathcal{D}_2 = 0$ on I.

<u>Case 2</u>: $(w_5, w_6) = (0, 0)$. This implies $w_4 \neq 0$. On these assumptions

$$\mathcal{D}_3 = (w_1 w_3 - w_2^2)(1 + z_4 w_4) + w_3 w_4$$

and

$$\gamma^3(\mathcal{D}_3) = (w_1 w_3 - w_2^2) w_4.$$

Now $\gamma^3(\mathcal{D}_3) = 0$ implies that $w_1 w_3 - w_2^2 = 0$, which in combination with $\mathcal{D}_3 = 0$ implies that $w_3 = 0$. This in turn gives $w_2 = 0$ and

$$\begin{pmatrix} w_1 & w_2 & w_4 & w_5 \\ w_2 & w_3 & w_5 & w_6 \end{pmatrix} = \begin{pmatrix} w_1 & 0 & w_4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

which contradicts the assumption that

$$\operatorname{Rank} \begin{pmatrix} w_1 & w_2 & w_4 & w_5 \\ w_2 & w_3 & w_5 & w_6 \end{pmatrix} = 2.$$

Finally we study the stratum of *nongeneric fibers*, that is, $(a_3, a_4) = (0, 0)$. Here all fibers are smooth. Also

$$\mathcal{F}^5_{(a_1,a_2,0,0)} = \mathcal{F}^4_{(a_1,a_2,0,0)} \times \mathbb{C}^3$$

and since $\mathcal{F}^4_{(a_1,a_2,0,0)}$ is elliptic we are done.

14. Proof of Proposition 3.6: induction steps

Recall the description of the stratification for the submersion $\Phi_M = \pi_4 \circ \Psi_M$ given in Section 7. When M is *odd* we have the following strata:

- The strata of *generic fibers*: When $(a_3, a_4) \neq (0, 0)$, the fibers are graphs over $\mathcal{G}^{M-1}_{(a_3,a_4)} \times \mathbb{C}$. This set is divided into two strata as follows:
  - Smooth generic fibers: When $(a_3, a_4) \neq (0, 1)$, the fibers are smooth.
  - Singular generic fibers: When $(a_3, a_4) = (0, 1)$, the fibers are nonsmooth.
- The stratum of *nongeneric fibers*: When $(a_3, a_4) = (0, 0)$ the fibers are $\mathcal{F}^M_{(a_1,a_2,0,0)} = \mathcal{F}^{M-1}_{(a_1,a_2,0,0)} \times \mathbb{C}^3$. Moreover the fibers are smooth.

When M is *even* we have the following strata:

- The stratum of *generic fibers*: When $(a_1, a_2) \neq (0, 0)$, the fibers are graphs over $\mathcal{H}^{M-1}_{(a_1,a_2)} \times \mathbb{C}$. Moreover the fibers are smooth.
- The strata of *nongeneric fibers*: When $(a_1, a_2) = (0, 0)$, the fibers are $\mathcal{F}^M_{(0,0,a_3,a_4)} = \mathcal{F}^{M-1}_{(0,0,a_3,a_4)} \times \mathbb{C}^3$. This set is divided into two strata as follows:
  - Smooth nongeneric fibers: When $(a_3, a_4) \neq (0, 1)$, the fibers are smooth.
  - Singular nongeneric fibers: When $(a_3, a_4) = (0, 1)$, the fibers are nonsmooth.

We will now complete the proof by doing the induction steps necessary.

14.1. ***Even number of factors.*** We begin by showing that the stratified submersion is elliptic when the number of matrix factors is even. This case is easier than the case when the number of factors is odd, which we will deal with in Section 14.2. Assume that $K = 2k-1 \geq 5$ and that the submersions $\Phi_L = \pi_4 \circ \Psi_L$ are stratified elliptic submersions when $3 \leq L \leq K$ and that Proposition 10.9 is true when $3 \leq L \leq K$.

We study

$$\mathcal{F}^{K+1}_{(a_1,a_2,a_3,a_4)} = \mathcal{F}^{2k}_{(a_1,a_2,a_3,a_4)} = \bigcup_{(w_{3k-2},w_{3k-1},w_{3k})\in\mathbb{C}^3} \mathcal{F}^{2k-1}_{(a_1,a_2,b_3,b_4)}, \tag{14.1.1}$$

where $b_3 = a_3 - w_{3k-2}a_1 - w_{3k-1}a_2$ and $b_4 = a_4 - w_{3k-1}a_1 - w_{3k}a_2$. That is, we use the new group of variables w_{3k-2}, w_{3k-1} and w_{3k} to present $\mathcal{F}^{2k}_{(a_1,a_2,a_3,a_4)}$ as a fibration over $\mathbb{C}^3$ with fibers $\mathcal{F}^{2k-1}$.

Let us describe the strategy similar to the cases of four and five matrix factors. We want to use Proposition 10.9 for $K = 2k-1$, which gives us complete fields that span along that fibration. Next we want to find complete fields among those that are tangential to $\mathcal{F}^{2k}_{(a_1,a_2,a_3,a_4)}$ that also are transversal to the fibers in the fibration. We will appeal to Corollary 10.4 and Lemma 10.5(1) to find these fields. Taken together this will show that a subset A in the fiber $\mathcal{F}^{2k}_{(a_1,a_2,a_3,a_4)}$ that is invariant with respect to vector fields from $\mathcal{Q}_{2k}$ must be contained in the union of nongeneric fibers $\mathcal{F}^{2k-1}$ and singular points of generic fibers $\mathcal{F}^{2k-1}$. Call this union $\mathcal{U}^{2k}_{(a_1,a_2,a_3,a_4)}$. Our aim will then be to show that there cannot exist such an invariant set A by showing that every point in $\mathcal{U}^{2k}_{(a_1,a_2,a_3,a_4)}$ can be moved into $\mathcal{F}^{2k}_{(a_1,a_2,a_3,a_4)} \setminus \mathcal{U}^{2k}_{(a_1,a_2,a_3,a_4)}$ by vector fields in $\mathcal{Q}_{2k}$.

We now take care of the details. By (14.1.1) there are b_3 and b_4 so that $\vec{Z}_{2k-1} \in \mathcal{F}^{2k-1}_{(a_1,a_2,b_3,b_4)}$ and

$$\vec{Z}_{2k} = (\vec{Z}_{2k-1}, w_{3k-2}, w_{3k-1}, w_{3k}).$$

We begin by studying the stratum of *generic fibers*, that is, $(a_1,a_2) \neq (0,0)$. For points where $(b_3,b_4) \notin \{(0,0),(0,1)\}$ we have that $\mathcal{F}^{2k-1}_{(a_1,a_2,b_3,b_4)}$ is a smooth generic fiber for the submersion Φ_{2k-1}, and Proposition 10.9 (for $L = 2k-1$) together with Corollary 10.4 and Lemma 10.5(1) let us conclude that we have complete vector fields spanning the tangent space of $\mathcal{F}^{2k}_{(a_1,a_2,a_3,a_4)}$ at these points. For points where $(b_3,b_4) = (0,0)$ we have

$$\mathcal{F}^{2k-1}_{(a_1,a_2,0,0)} = \mathcal{F}^{2k-2}_{(a_1,a_2,0,0)} \times \mathbb{C}^3$$

and Proposition 10.9 (for $L = 2k-2$ applied to the first factor) together with Corollary 10.4 and Lemma 10.5(1) (Lemma 10.5(2) when $2k-2 = 4$) show that we have spanning fields in these points. If $(b_3,b_4) = (0,1)$ then $\mathcal{F}^{2k-1}_{(a_1,a_2,0,1)}$ is a singular generic fiber for Φ_{2k-1} and at smooth points of the fiber we have complete spanning fields by Proposition 10.9 (for $L = 2k-1$), Corollary 10.4 and Lemma 10.5. It remains to study

$$\vec{Z}_{2k-1} \in \mathrm{Sing}(\mathcal{F}^{2k-1}_{(a_1,a_2,0,1)}),$$

which is given by

$$z_2 = z_3 = z_5 = z_6 = \cdots = z_{3k-4} = z_{3k-3} = 0 \tag{14.1.2}$$

and

$$\mathrm{Rank}\begin{pmatrix} w_1 & w_2 & \ldots & w_{3k-5} & w_{3k-4} \\ w_2 & w_3 & \ldots & w_{3k-4} & w_{3k-3} \end{pmatrix} < 2. \tag{14.1.3}$$

A calculation assuming (14.1.2) shows that

$$\partial^{2k-1}_{z_{3k-4}z_{3k-3}z_{3k}} = (z_{3k-2}w_{3k-4} + z_{3k-1}w_{3k-3})\frac{\partial}{\partial z_{3k-4}} - (1 + z_{3k-2}w_{3k-5} + z_{3k-1}w_{3k-4})\frac{\partial}{\partial z_{3k-3}} + \cdots .$$

Therefore the complete fields

$$\begin{aligned}\theta^{2k}_{z_{3k-4}z_{3k-3}z_{3k}} &= a_1^2\partial^{2k-1}_{z_{3k-4}z_{3k-3}z_{3k}} + \cdots ,\\ \phi^{2k}_{z_{3k-4}z_{3k-3}z_{3k}} &= a_2^2\partial^{2k-1}_{z_{3k-4}z_{3k-3}z_{3k}} + \cdots\end{aligned}$$

move points out of $\mathrm{Sing}(\mathcal{F}^{2k-1}_{(a_1,a_2,0,1)})$ (into the big orbit) unless in addition to (14.1.2) and (14.1.3) also

$$z_{3k-2}w_{3k-4} + z_{3k-1}w_{3k-3} = 1 + z_{3k-2}w_{3k-5} + z_{3k-1}w_{3k-4} = 0. \tag{14.1.4}$$

Points in an invariant subset must satisfy also these equations. A calculation assuming (14.1.2) gives that $\partial^{2k-1}_{z_{3k-2}z_{3k-1}z_{3k}} = \partial/\partial z_{3k-2}$. Therefore the complete fields

$$\begin{aligned}\theta^{2k}_{z_{3k-2}z_{3k-1}z_{3k}} &= a_1^2\partial^{2k-1}_{z_{3k-2}z_{3k-1}z_{3k}} + \cdots ,\\ \theta^{2k}_{z_{3k-2}z_{3k-1}z_{3k}} &= a_2^2\partial^{2k-1}_{z_{3k-2}z_{3k-1}z_{3k}} + \cdots\end{aligned}$$

move points out of this set since

$$\begin{aligned}\theta^{2k}_{z_{3k-2}z_{3k-1}z_{3k}}(1 + z_{3k-2}w_{3k-5} + z_{3k-1}w_{3k-4}) &= a_1^2 w_{3k-5},\\ \theta^{2k}_{z_{3k-2}z_{3k-1}z_{3k}}(z_{3k-2}w_{3k-4} + z_{3k-1}w_{3k-3}) &= a_1^2 w_{3k-4},\\ \phi^{2k}_{z_{3k-2}z_{3k-1}z_{3k}}(1 + z_{3k-2}w_{3k-5} + z_{3k-1}w_{3k-4}) &= a_2^2 w_{3k-5},\\ \phi^{2k}_{z_{3k-2}z_{3k-1}z_{3k}}(z_{3k-2}w_{3k-4} + z_{3k-1}w_{3k-3}) &= a_2^2 w_{3k-4}\end{aligned}$$

cannot all be zero, because this would contradict (14.1.4). Notice that this proves Proposition 10.9 for $L = 2k$.

Now we study the stratum of *nongeneric fibers*, that is, $a_1 = a_2 = 0$. In this case we know that

$$\mathcal{F}^{2k}_{(0,0,a_3,a_4)} = \mathcal{F}^{2k-1}_{(0,0,a_3,a_4)} \times \mathbb{C}^3$$

and by the induction assumption we are done. This finishes the induction step for an even number of factors.

14.2. ***Odd number of factors.*** We assume that $K = 2k \geq 6$ and that the submersions $\Phi_L = \pi_4 \circ \Psi_L$ are stratified elliptic submersions when $3 \leq L \leq K$ and that Proposition 10.9 is true when $3 \leq L \leq K$.

We study

$$\mathcal{F}^{K+1}_{(a_1,a_2,a_3,a_4)} = \mathcal{F}^{2k+1}_{(a_1,a_2,a_3,a_4)} = \bigcup_{(z_{3k+1},z_{3k+2},z_{3k+3})\in\mathbb{C}^3} \mathcal{F}^{2k}_{(b_1,b_2,a_3,a_4)}, \tag{14.2.1}$$

where $b_1 = a_1 - z_{3k+1}a_3 - z_{3k+2}a_4$ and $b_2 = a_2 - z_{3k+2}a_3 - z_{3k+3}a_4$. Let $\vec{Z}_{2k+1} \in \mathcal{F}^{2k+1}_{(a_1,a_2,a_3,a_4)}$. Because of (14.2.1) there are b_1 and b_2 so that $\vec{Z}_{2k} \in \mathcal{F}^{2k}_{(b_1,b_2,a_3,a_4)}$ and

$$\vec{Z}_{2k+1} = (\vec{Z}_{2k}, z_{3k+1}, z_{3k+2}, z_{3k+3}).$$

Begin with the stratum of *smooth generic fibers*, that is,

$$(a_3, a_4) \notin \{(0,0), (0,1)\}.$$

First notice that if $(b_1, b_2) \neq (0,0)$ then $\mathcal{F}^{2k}_{(b_1,b_2,a_3,a_4)}$ is a generic smooth fiber for Φ_{2k} and as above, Proposition 10.9 (for $L = 2k$), Corollary 10.4 and Lemma 10.5 show that for these points we have spanning fields. If $(b_1, b_2) = (0,0)$ then

$$\mathcal{F}^{2k}_{(0,0,a_3,a_4)} \cong \mathbb{C}^3 \times \mathcal{F}^{2k-1}_{(0,0,a_3,a_4)}$$

is a nongeneric smooth fiber for Φ_{2k} and, since $\mathcal{F}^{2k-1}_{(0,0,a_3,a_4)}$ is a generic smooth fiber, Proposition 10.9 (for $L = 2k-1$), Corollary 10.4 and Lemma 10.5 show that for these points we have spanning fields.

We now study the stratum of *singular generic fibers*. Here $(a_3, a_4) = (0,1)$. Again notice that if $(b_1, b_2) \neq (0,0)$ then $\mathcal{F}^{2k}_{(b_1,b_2,0,1)}$ is a generic smooth fiber for Φ_{2k}, and Proposition 10.9 (for $L = 2k$), Corollary 10.4 and Lemma 10.5 show that for these points we have spanning fields as above. Next we study the case $(b_1, b_2) = (0,0)$. In this case we see that $\mathcal{F}^{2k}_{(0,0,0,1)}$ is a singular nongeneric fiber of Φ_{2k} and

$$\mathcal{F}^{2k}_{(0,0,0,1)} \cong \mathcal{F}^{2k-1}_{(0,0,0,1)} \times \mathbb{C}^3_{w_{3k-2}w_{3k-1}w_{3k}}.$$

The smooth points of $\mathcal{F}^{2k-1}_{(0,0,0,1)}$ (which is generic) are handled using Proposition 10.9 (for $L = 2k-1$), Corollary 10.4 and Lemma 10.5. We have the chain of inclusions

$$\mathcal{F}^{2k-1}_{(a_1,a_2,0,1)} \supset \mathcal{F}^{2k}_{(0,0,0,1)} = \mathcal{F}^{2k-1}_{(0,0,0,1)} \times \mathbb{C}^3 \supset \mathrm{Sing}(\mathcal{F}^{2k}_{(0,0,0,1)}) \times \mathbb{C}^3 \supset \mathrm{Sing}(\mathcal{F}^{2k+1}_{(a_1,a_2,0,1)}).$$

By the arguments above any possible invariant subset must be contained in

$$J = (\mathrm{Sing}(\mathcal{F}^{2k}_{(0,0,0,1)}) \times \mathbb{C}^3) \setminus \mathrm{Sing}(\mathcal{F}^{2k+1}_{(a_1,a_2,0,1)}).$$

Points in J are characterized by $z_2 = z_3 = \cdots = z_{3k-4} = z_{3k-3} = z_{3k-1} = z_{3k} = 0$,

$$\mathrm{Rank}\begin{pmatrix} w_1 & w_2 & \dots & w_{3k-5} & w_{3k-4} \\ w_2 & w_3 & \dots & w_{3k-4} & w_{3k-3} \end{pmatrix} < 2$$

and

$$\mathrm{Rank}\begin{pmatrix} w_1 & w_2 & \dots & w_{3k-5} & w_{3k-4} & w_{3k-2} & w_{3k-1} \\ w_2 & w_3 & \dots & w_{3k-4} & w_{3k-3} & w_{3k-1} & w_{3k} \end{pmatrix} = 2.$$

Take the largest $l < k$ such that

$$\mathrm{Rank}\begin{pmatrix} w_{3l-2} & w_{3l-1} \\ w_{3l-1} & w_{3l} \end{pmatrix} = 1.$$

Let $\widehat{Z} = \sum_{j=l+1}^{k} z_{3j-2}$. We examine the complete field $\phi^{2k+1}_{z_{3l-1}z_{3l}z_{3k}}$. This field has some complicated components which on J take the form

$$\phi^{2k+1}_{z_{3l-1}z_{3l}z_{3k}} = \mathcal{D}_1 \frac{\partial}{\partial z_{3l-1}} + \mathcal{D}_2 \frac{\partial}{\partial z_{3l}} + \mathcal{D}_3 \frac{\partial}{\partial z_{3k}} + \cdots,$$

where

$$\mathcal{D}_1 = \det\begin{pmatrix} w_{3l-1}+w_{3k-1}+w_{3l-1}w_{3k-2}\widehat{Z} & w_{3k-1} \\ w_{3l}+w_{3k}+w_{3l-1}w_{3k-1}\widehat{Z} & w_{3k}\end{pmatrix},$$

$$\mathcal{D}_2 = \det\begin{pmatrix} w_{3l-2}+w_{3k-2}+w_{3l-2}w_{3k-2}\widehat{Z} & w_{3k-1} \\ w_{3l-1}+w_{3k-1}+w_{3l-2}w_{3k-1}\widehat{Z} & w_{3k}\end{pmatrix},$$

$$\mathcal{D}_3 = \det\begin{pmatrix} w_{3l-2}+w_{3k-2}+w_{3l-2}w_{3k-2}\widehat{Z} & w_{3l-1}+w_{3k-1}+w_{3l-1}w_{3k-2}\widehat{Z} \\ w_{3l-1}+w_{3k-1}+w_{3l-2}w_{3k-1}\widehat{Z} & w_{3l}+w_{3k}+w_{3l-1}w_{3k-1}\widehat{Z}\end{pmatrix}.$$

Whenever at least one of $\mathcal{D}_1$, $\mathcal{D}_2$ or $\mathcal{D}_3$ is nonzero we can move out of J. Now suppose we are in a point of J where $\mathcal{D}_1 = \mathcal{D}_2 = \mathcal{D}_3 = 0$.

Let

$$\mathcal{C} = \begin{pmatrix} w_{3l-2}+w_{3k-2}+w_{3l-2}w_{3k-2}\widehat{Z} & w_{3l-1}+w_{3k-1}+w_{3l-1}w_{3k-2}\widehat{Z} & w_{3k-2} & w_{3k-1} \\ w_{3l-1}+w_{3k-1}+w_{3l-2}w_{3k-1}\widehat{Z} & w_{3l}+w_{3k}+w_{3l-1}w_{3k-1}\widehat{Z} & w_{3k-1} & w_{3k}\end{pmatrix}$$

and observe that

$$2 = \operatorname{Rank}\begin{pmatrix} w_{3l-2} & w_{3l-1} & w_{3k-2} & w_{3k-1} \\ w_{3l-1} & w_{3l} & w_{3k-1} & w_{3k}\end{pmatrix} = \operatorname{Rank}\mathcal{C}$$

by column operations.

The fact that $\mathcal{D}_1 = \mathcal{D}_2 = \mathcal{D}_3 = 0$ means that the rank drops when we remove the third column from these matrices. This implies that the third column is nonzero and the other columns are multiples of a nonzero vector v which, moreover, is linearly independent of the third column. Now we use the field γ^{3l} (see (8.0.13) or (8.0.17)) to show that the set

$$I = J \cap \{\mathcal{D}_1 = \mathcal{D}_2 = \mathcal{D}_3 = 0\}$$

does not contain an invariant subset under fields from $\mathcal{Q}_{2k+1}$. In the points that we are considering, $\gamma^{3l} = \partial/\partial z_{3l+1}$. We consider two cases.

Case 1: $(w_{3k-1}, w_{3k}) \neq (0, 0)$. In this case

$$\det\begin{pmatrix} w_{3k-2} & w_{3k-1} \\ w_{3k-1} & w_{3k}\end{pmatrix} \neq 0.$$

We have

$$\gamma^{3l}(\mathcal{D}_1) = w_{3l-1}\det\begin{pmatrix} w_{3k-2} & w_{3k-1} \\ w_{3k-1} & w_{3k}\end{pmatrix}.$$

Thus γ^{3l} moves points out of I unless $w_{3l-1} = 0$. Looking at

$$\gamma^{3l}(\mathcal{D}_2) = w_{3l-2}\det\begin{pmatrix} w_{3k-2} & w_{3k-1} \\ w_{3k-1} & w_{3k}\end{pmatrix},$$

we see that $w_{3l-2} = 0$ for I to be invariant. Assuming in addition $w_{3l-2} = w_{3l-1} = 0$ we find that

$$\mathcal{D}_2 = \det\begin{pmatrix} w_{3k-2} & w_{3k-1} \\ w_{3k-1} & w_{3k}\end{pmatrix} = 0,$$

which is a contradiction.

Case 2: $(w_{3k-1}, w_{3k}) = (0,0)$. This implies $w_{3k-2} \neq 0$. By these assumptions

$$\mathcal{D}_3 = (w_{3l-2}w_{3l} - w_{3l-1}^2)(1 + w_{3k-2}\widehat{Z}) + w_{3l}w_{3k-2}$$

and

$$\gamma^{3l}(\mathcal{D}_3) = (w_{3l-2}w_{3l} - w_{3l-1}^2)w_{3k-2}.$$

Now $\mathcal{D}_3 = \gamma^{3l}(\mathcal{D}_3) = 0$ implies that $(w_{3l-2}w_{3l} - w_{3l-1}^2) = 0$ and $w_{3l} = 0$. The first equality gives $w_{3l-1} = 0$, which altogether contradicts the assumption that

$$\operatorname{Rank}\begin{pmatrix} w_{3l-2} & w_{3l-1} & w_{3k-2} & w_{3k-1} \\ w_{3l-1} & w_{3l} & w_{3k-1} & w_{3k} \end{pmatrix} = 2.$$

Finally we study the stratum of *nongeneric fibers*, that is, $(a_3, a_4) = (0,0)$. Here all fibers are smooth. Also

$$\mathcal{F}^{2k+1}_{(a_1,a_2,0,0)} = \mathcal{F}^{2k}_{(a_1,a_2,0,0)} \times \mathbb{C}^3$$

and since $\mathcal{F}^{2k}_{(a_1,a_2,0,0)}$ is elliptic, by the induction hypothesis we are done.

15. Product of exponentials and open questions

For a Stein space X, a complex Lie group G and its exponential map $\exp : \mathfrak{g} \to G$, we say that a holomorphic map $f : X \to G$ is a product of k exponentials if there are holomorphic maps $f_1, \dots, f_k : X \to \mathfrak{g}$ such that

$$f = \exp(f_1) \cdots \exp(f_k).$$

It is easy to see that any map f which is a product of exponentials (for some sufficiently large k) is null-homotopic. In the case where G is the special linear group $\mathrm{SL}_n(\mathbb{C})$ the converse follows from [Ivarsson and Kutzschebauch 2012] as explained in [Doubtsov and Kutzschebauch 2019]. In the same way we prove:

Theorem 15.1. *For a Stein space X there is a number N depending on the dimension of X such that any null-homotopic holomorphic map $f : X \to \mathrm{Sp}_4(\mathbb{C})$ can be factorized as*

$$f(x) = \exp(G_1(x)) \cdots \exp(G_K(x)).$$

where $G_i : X \to \mathfrak{sp}_4(\mathbb{C})$ are holomorphic maps.

Proof. By Theorem 3.1 we find K elementary symplectic matrices $A_i(x) \in \mathrm{Sp}_4(\mathcal{O}(X))$, $i = 1, 2, \dots K$, such that

$$f(x) = A_1(x) \cdots A_K(x).$$

Now remark that the logarithmic series

$$\ln(\mathrm{Id} + B) = \sum \frac{1}{n} B^n$$

is finite for the nilpotent matrices $B_i = A_i - \mathrm{Id}$. □

Open Problem 15.2. *Determine the optimal number K in Theorem 15.1.*

Open Problem 15.3. *Determine the optimal numbers of factors in Theorem 3.1.*

The smooth fibers

$$\mathcal{F}^K_{(a_1,a_2,a_3,a_4)} = (\pi_4 \circ \Psi_K)^{-1}(a_1, a_2, a_3, a_4)$$

of the fibration projecting the product of K elementary symplectic matrices to its last row are smooth affine algebraic varieties. They are new examples of Oka manifolds, since we prove as a by-product of Proposition 3.6 that they are holomorphically flexible (for a definition see the work of Arzhantsev, Flenner, Kaliman, Kutzschebauch and Zaidenberg [Arzhantsev et al. 2013]). Our proof does not give the algebraic flexibility of them, even if our initial complete fields obtained in Section 8 are algebraic. The problem is that their flows are not always algebraic (not all of them are locally nilpotent). Therefore the pull-backs by their flows are merely holomorphic vector fields.

Open Problem 15.4. *Which other* (*stronger*) *flexibility properties like algebraic flexibility*, *algebraic* (*volume*) *density property*, *or* (*volume*) *density property do the fibers* $\mathcal{F}^K_{(a_1,a_2,a_3,a_4)}$ *admit*?

For the definition of these flexibility properties we refer to the overview article [Kutzschebauch 2014].

Let us remark that the fibers of the fibration for five elementary factors in [Ivarsson and Kutzschebauch 2012] have been thoroughly studied in [Kaliman and Kutzschebauch 2011; 2016, Section 7]. They were the starting point for the introduction of the class of generalized Gizatullin surfaces whose final classification was achieved by Kaliman, Kutzschebauch and Leuenberger [Kaliman et al. 2020]. The topology of these fibers for any number of elementary factors has been studied in [De Vito 2020], where it was also proven that they admit the algebraic volume density property. Such studies are interesting since the possible topological types of Oka manifolds or manifolds with the density property are not understood at the moment.

Open Problem 15.5. *Determine the homology groups of the fibers* $\mathcal{F}^K_{(a_1,a_2,a_3,a_4)}$.

And finally:

Open Problem 15.6. *Prove Conjecture 3.11.*

References

[Andrist and Kutzschebauch 2018] R. B. Andrist and F. Kutzschebauch, "The fibred density property and the automorphism group of the spectral ball", *Math. Ann.* **370**:1-2 (2018), 917–936. MR Zbl

[Arzhantsev et al. 2013] I. Arzhantsev, H. Flenner, S. Kaliman, F. Kutzschebauch, and M. Zaidenberg, "Flexible varieties and automorphism groups", *Duke Math. J.* **162**:4 (2013), 767–823. MR Zbl

[Bass et al. 1967] H. Bass, J. Milnor, and J.-P. Serre, "Solution of the congruence subgroup problem for SL_n $(n \geq 3)$ and Sp_{2n} $(n \geq 2)$", *Inst. Hautes Études Sci. Publ. Math.* **33** (1967), 59–137. MR Zbl

[Cohn 1966] P. M. Cohn, "On the structure of the GL_2 of a ring", *Inst. Hautes Études Sci. Publ. Math.* **30** (1966), 5–53. MR Zbl

[De Vito 2020] G. De Vito, "New examples of Stein manifolds with volume density property", *Complex Anal. Synerg.* **6**:2 (2020), art. id. 9. MR Zbl

[Doubtsov and Kutzschebauch 2019] E. Doubtsov and F. Kutzschebauch, "Factorization by elementary matrices, null-homotopy and products of exponentials for invertible matrices over rings", *Anal. Math. Phys.* **9**:3 (2019), 1005–1018. MR Zbl

[Draisma 2022] J. Draisma, "The irreducible control property in matrix groups", *Linear Algebra Appl.* **634** (2022), 15–29. MR

[Forstnerič 2010] F. Forstnerič, "The Oka principle for sections of stratified fiber bundles", *Pure Appl. Math. Q.* **6**:3 (2010), 843–874. MR Zbl

[Forstnerič 2017] F. Forstnerič, *Stein manifolds and holomorphic mappings*: *the homotopy principle in complex analysis*, 2nd ed., Ergebnisse der Mathematik (3) **56**, Springer, 2017. MR Zbl

[Forstnerič and Prezelj 2001] F. Forstnerič and J. Prezelj, "Extending holomorphic sections from complex subvarieties", *Math. Z.* **236**:1 (2001), 43–68. MR Zbl

[Forstnerič and Prezelj 2002] F. Forstnerič and J. Prezelj, "Oka's principle for holomorphic submersions with sprays", *Math. Ann.* **322**:4 (2002), 633–666. MR Zbl

[Gromov 1989] M. Gromov, "Oka's principle for holomorphic sections of elliptic bundles", *J. Amer. Math. Soc.* **2**:4 (1989), 851–897. MR Zbl

[Grunewald et al. 1991] F. Grunewald, J. Mennicke, and L. Vaserstein, "On symplectic groups over polynomial rings", *Math. Z.* **206**:1 (1991), 35–56. MR Zbl

[Ivarsson and Kutzschebauch 2012] B. Ivarsson and F. Kutzschebauch, "Holomorphic factorization of mappings into $\mathrm{SL}_n(\mathbb{C})$", *Ann. of Math.* (2) **175**:1 (2012), 45–69. MR Zbl

[Ivarsson et al. 2020] B. Ivarsson, F. Kutzschebauch, and E. Løw, "Factorization of symplectic matrices into elementary factors", *Proc. Amer. Math. Soc.* **148**:5 (2020), 1963–1970. MR Zbl

[Kaliman and Kutzschebauch 2011] S. Kaliman and F. Kutzschebauch, "On the present state of the Andersén–Lempert theory", pp. 85–122 in *Affine algebraic geometry* (Montreal, 2009), edited by D. Daigle et al., CRM Proc. Lect. Notes **54**, Amer. Math. Soc., Providence, RI, 2011. MR Zbl

[Kaliman and Kutzschebauch 2016] S. Kaliman and F. Kutzschebauch, "On algebraic volume density property", *Transform. Groups* **21**:2 (2016), 451–478. MR Zbl

[Kaliman et al. 2020] S. Kaliman, F. Kutzschebauch, and M. Leuenberger, "Complete algebraic vector fields on affine surfaces", *Int. J. Math.* **31**:3 (2020), art. id. 2050018. MR Zbl

[Klein and Ramspott 1987] M. Klein and K. J. Ramspott, "Ein Transformationssatz für Idealbasen holomorpher Funktionen", *Bayer. Akad. Wiss. Math.-Natur. Kl. Sitzungsber.* **1987**:4 (1987), 93–100. MR Zbl

[Kopeiko 1978] V. I. Kopeiko, "The stabilization of symplectic groups over a polynomial ring", *Mat. Sb.* (*N.S.*) **106(148)**:1 (1978), 94–107. In Russian; translated in *Math. USSR-Sb.* **34**:5 (1978), 655–669. MR Zbl

[Kutzschebauch 2014] F. Kutzschebauch, "Flexibility properties in complex analysis and affine algebraic geometry", pp. 387–405 in *Automorphisms in birational and affine geometry* (Levico Terme, Italy, 2012), edited by I. Cheltsov et al., Springer Proc. Math. Stat. **79**, Springer, 2014. MR Zbl

[Suslin 1977] A. A. Suslin, "The structure of the special linear group over rings of polynomials", *Izv. Akad. Nauk SSSR Ser. Mat.* **41**:2 (1977), 235–252. In Russian; translated in *Math. USSR-Izv.* **11**:2 (1977), 221–238. MR Zbl

[Vaserstein 1988] L. N. Vaserstein, "Reduction of a matrix depending on parameters to a diagonal form by addition operations", *Proc. Amer. Math. Soc.* **103**:3 (1988), 741–746. MR Zbl

[Vavilov and Stepanov 2011] N. A. Vavilov and A. V. Stepanov, "Linear groups over general rings, I: Generalities", *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* (*POMI*) **394**:22 (2011), 33–139. In Russian; translated in *J. Math. Sci.* (*N.Y.*) **188**:5 (2013), 490–550. MR Zbl

Received 28 May 2020. Revised 8 May 2021. Accepted 10 Jun 2021.

BJÖRN IVARSSON: bjorn.ivarsson@aalto.fi
Department of Mathematics and Systems Analysis, Aalto University, Aalto, Finland

FRANK KUTZSCHEBAUCH: frank.kutzschebauch@math.unibe.ch
Departement Mathematik, Universität Bern, Bern, Switzerland

ERIK LØW: elow@math.uio.no
Department of Mathematics, University of Oslo, Oslo, Norway

mathematical sciences publishers msp

ANALYSIS AND PDE
Vol. 16 (2023), No. 1, pp. 279–308
DOI: 10.2140/apde.2023.16.279

TRANSVERSAL FAMILIES OF NONLINEAR PROJECTIONS AND GENERALIZATIONS OF FAVARD LENGTH

ROSEMARIE BONGERS AND KRYSTAL TAYLOR

Projections detect information about the size, geometric arrangement, and dimension of sets. To approach this, one can study the energies of measures supported on a set and the energies for the corresponding pushforward measures on the projection side. For orthogonal projections, quantitative estimates rely on a separation condition: most points are well-differentiated by most projections. It turns out that this idea also applies to a broad class of nonlinear projection-type operators satisfying a *transversality condition*. We establish that several important classes of nonlinear projections are transversal. This leads to quantitative lower bounds for decay rates for nonlinear variants of Favard length, including Favard curve length (as well as a new generalization to higher dimensions, called Favard surface length) and visibility measurements associated to radial projections. As one application, we provide a simplified proof for the decay rate of the Favard curve length of generations of the four-corner Cantor set, first established by Cladek, Davey, and Taylor.

1. Introduction and main results

The Favard length of a planar set E is the average length of its orthogonal projections. It is defined by

$$\operatorname{Fav}(E) = \frac{1}{\pi}\int_0^\pi |P_\theta(E)|\, d\theta,$$

where P_θ is orthogonal projection into a line L_θ through the origin at angle θ from the positive x-axis and $|\cdot|$ denotes the 1-dimensional Hausdorff measure. Favard length gives a 1-dimensional notion of the *size* of a set which takes into account the geometry, arrangement, and rectifiability of the underlying set. As a consequence, there are deep relationships between Favard length and analytic capacity, the understanding of which is related to important open problems in geometric measure theory. As we will see, variants of the Favard length can also be formulated for more general families of mappings, beyond the orthogonal projections, and in higher dimensions.

As the Hausdorff dimension of a set cannot increase under a projection, sets of dimension $s < 1$ have Favard length equal to zero. A refinement due to Marstrand [1954] actually shows that the dimension of such a set will be preserved in almost every direction. On the other hand, sets with dimension $s > 1$ will have positive-length projections in almost every direction, and therefore have positive Favard length. Therefore, the critical dimension is $s = 1$.

Taylor is supported in part by the Simons Foundation grant 523555.
MSC2020: 28A75, 28A80, 57N75.
Keywords: nonlinear projections, transversality, Favard length, fractals, transversal projections, Buffon needle, energy, rate of decay, radial projections, curve projections, rectifiability.

In dimension 1, the key geometric property that Favard length can detect is *rectifiability*: it is a consequence of the Besicovitch projection theorem [1939] that purely unrectifiable sets in the plane with finite 1-dimensional Hausdorff measure have Favard length equal to zero. (For an exposition of the full Besicovitch–Federer projection theorem in all dimensions, see [Mattila 1995, Chapter 18].) While Besicovitch's theorem gives a qualitative result, we can find related quantitative theorems. If $E(r)$ is the r-neighborhood of a set E with Favard length zero, the dominated convergence theorem shows that

$$\lim_{r\to 0^+} \mathrm{Fav}(E(r)) = 0.$$

More precise asymptotic information for $\mathrm{Fav}(E(r))$ as r decreases to zero can give quantitative measurements of the dimension, size, and geometric arrangement of E. A number of authors have investigated quantitative versions of the Besicovitch projection theorem for general sets. The best known results in terms of upper and lower bounds are due to Tao [2009] and Mattila [1990] respectively.

Tao introduced a quantitative version of rectifiability for sets in the plane of finite $\mathcal{H}^1$ measure and used multiscale analysis to show that an upper bound on the so-called rectifiability constant yields an upper bound on the Favard length. A nonlinear version of Tao's theorem is studied in a work of Davey and the second author [Davey and Taylor 2022].

Mattila [1990] established a fundamental relationship between the Favard length of a set and its Hausdorff dimension. In two dimensions, it states:

Theorem 1.1 (Favard lengths for neighborhoods [Mattila 1990]). *Fix $s \in (0, 1]$. If $F \subseteq \mathbb{R}^2$ is the support of a Borel probability measure with $\mu(B(x, r)) \le br^s$ for all $x \in \mathbb{R}^2$ and $0 < r < \infty$, then*

$$\mathrm{Fav}(F(r)) \gtrsim r^{1-s}$$

if $s < 1$ and

$$\mathrm{Fav}(F(r)) \gtrsim (\log r^{-1})^{-1}$$

if $s = 1$.

Throughout the paper, we will use the notation $A \lesssim B$ to mean that there is a constant C so that $A \le CB$ and will write $A \sim B$ if $A \lesssim B$ and $B \lesssim A$.

The proof of Mattila's result follows from studying *energies*: if μ is a measure, its *s-energy* is

$$I_s(\mu) = \iint \frac{d\mu(x)\, d\mu(y)}{|x-y|^s}. \tag{1-1}$$

This quantity is closely tied to Hausdorff dimension; see, e.g., [Mattila 1995, Chapter 8] for a formulation of the definition of Hausdorff dimension in terms of s-energies. In order to relate a measure to the projections, we need the notion of a pushforward: if $f : X \to Y$ is a function and μ is a measure supported on X we will define the *pushforward measure* $f_\sharp\mu$ by

$$(f_\sharp\mu)(A) = \mu(f^{-1}(A)), \quad A \subseteq Y. \tag{1-2}$$

In general it can be difficult to study the pushforward of a particular mapping, yet it turns out that the *average* energy of a projection can be well controlled. That is, if $\{\pi_\alpha : \alpha \in A\}$ is an indexed family of orthogonal projections, it frequently is possible to precisely estimate

$$\int I_t(\pi_{\alpha\sharp}\mu)\,d\psi(\alpha),$$

where ψ is a measure on the index set A. By studying the average energy of the pushforwards of specialized measures supported on $F(r)$ with particular density properties, Mattila was able to establish the stated lower bounds. Further details are given in Section 3.

In the special setting that the underlying set is a fractal generated by an iterated function system, Mattila's techniques with energies are also applicable. A standard example of this is to consider the generations K_n of the *four-corner Cantor set*, which is defined by dividing the unit square into 16 axis-parallel squares of side length $\frac{1}{4}$, keeping the four-corner squares, and iterating the process within each corner. The limit of this process gives a prototypical example of a purely unrectifiable set with positive and finite length. As such, an important open problem is to estimate upper and lower bounds on the rate of decay in n of $\mathrm{Fav}(\mathcal{K}_n)$ (see [Łaba 2015] for a survey of results and techniques related to this problem). A variety of techniques can be used to show $\mathrm{Fav}(\mathcal{K}_n) \gtrsim n^{-1}$; see, for example, [Bongers 2019; Mattila 1990]. The tightest known results are

$$\frac{\log n}{n} \lesssim \mathrm{Fav}(\mathcal{K}_n) \lesssim \frac{1}{n^{1/6-\delta}} \tag{1-3}$$

for any $\delta > 0$, with the bounds due to Bateman and Volberg [2010] and Nazarov, Peres, and Volberg [Nazarov et al. 2010] respectively. Further, it is still a deep open question whether the Favard length $\mathrm{Fav}(\mathcal{K}_n)$ is larger or smaller than the analytic capacity $\gamma(\mathcal{K}_n)$, which is known to be of order $n^{-1/2}$ [Tolsa 2002].

The primary aim of this paper is to formulate Theorem 1.1 in a nonlinear setting for families of projections which are not orthogonal projections. In particular, we will consider families of maps satisfying the so-called transversality condition. After we establish a correspondence between the energy of a measure and its pushforwards under transversal families, we will apply these relationships to study the asymptotic decay rates of nonlinear variants of Favard length. In the process, we generalize the lower bounds on visibility established by Bond, Łaba, and Zahl [Bond et al. 2016], as well as provide a simplified proof of the lower bound for the Favard curve length of $\mathcal{K}_n$ derived by Cladek, Davey, and Taylor [Cladek et al. 2022]; both of these results are explored in Section 1A. Before stating our main results in Section 1C, we give several examples of families of nonlinear projection operators in Section 1A and we formalize the definition of transversality in Section 1B.

1A. ***Nonlinear projections.*** When orthogonal projections are replaced by more general families of nonlinear projection-type maps, one may ask if Besicovitch's theorem and its quantitative counterparts still hold. In many settings, these theorems still apply. Examples of such families include radial projections associated with visibility, curve-based projections associated with the Favard curve length and the surface projections we will introduce in this paper. Due to the special geometry exhibited by these projection families, the energy techniques of Mattila can be applied with appropriate modifications, leading to analogous lower bound on nonlinear Favard lengths.

1A1. *Visibility.* For a point $a \in \mathbb{R}^n$, the *radial projection* based at a maps $\mathbb{R}^n \setminus \{a\}$ to the $(n-1)$-dimensional unit sphere via

$$P_a(x) := \frac{x-a}{|x-a|}. \tag{1-4}$$

The *visibility* of a measurable set $E \subset \mathbb{R}^n$ from a *vantage point* a is

$$\operatorname{vis}(a, E) = |P_a(E)|, \tag{1-5}$$

where $|\cdot|$ denotes the $(n-1)$-dimensional Hausdorff measure on the unit sphere. In applications, we will restrict the vantage points a to a *vantage set* A. Informally, the visibility of a set E measures how much of the sky is filled up by the constellation E from an observer at vantage point a. As such, the set E is referred to as the *visible set*.

Bond, Łaba, and Zahl obtained upper and lower bounds on the visibility of δ-neighborhoods of unrectifiable self-similar 1-sets in the plane. In particular, their lower bound [Bond et al. 2016, Theorem 2.4] for visibility states that if μ is a positive, Borel, probability measure supported on a visible set $E \subset \mathbb{R}^2$ paired with an L-shaped vantage set $A \subset \mathbb{R}^2$ (with an extra separation condition), then

$$I_1(\mu)^{-1} \lesssim \int_A \operatorname{vis}(a, E)\, da.$$

Their work provides quantitative versions of the results in [Marstrand 1954; Simon and Solomyak 2006/07].

We will generalize this result by proving it for a wider range of vantage sets and extending it to higher dimensions. In particular, we provide a much weaker constraint on the geometric relationship between the vantage set and the visible set. As a particular application, we will demonstrate how such results can be used to obtain a lower bound on the rate of decay of the visibility of generations of the four-corner Cantor set from a wide variety of curves.

1A2. *Favard curve length.* As a second example of a context in which energy techniques can be applied, we define the family of maps which induce the *Favard curve length*. Let Γ denote a curve in $\mathbb{R}^2$. Given $\alpha \in \mathbb{R}$ and $(x, y) \in \mathbb{R}^2$, let $\Phi_\alpha(x, y)$ denote the set of y-coordinates of the intersection of $(x, y) + \Gamma$ with the line $\{x = \alpha\}$. That is,

$$\Phi_\alpha(x, y) = \{\beta \in \mathbb{R} : (\alpha, \beta) \in ((x, y) + \Gamma) \cap \{x = \alpha\}\}. \tag{1-6}$$

Given $\beta \in \mathbb{R}$, the inverse set $\Phi_\alpha^{-1}(\beta) = \{p \in \mathbb{R}^2 : \beta \in \Phi_\alpha(p)\}$ is given by $(\alpha, \beta) - \Gamma$. In the case that Γ can be expressed as the graph of a function and $\Phi_\alpha(x, y) \neq \varnothing$, the set $\Phi_\alpha(x, y)$ is a singleton and we identify $\Phi_\alpha(x, y)$ with that point.

If $E \subset \mathbb{R}^2$, then the *Favard curve length* of E is defined by

$$\operatorname{Fav}_\Gamma(E) := |\{(\alpha, \beta) \in \mathbb{R}^2 : \Phi_\alpha^{-1}(\beta) \cap E \neq \varnothing\}| = \int_{\mathbb{R}} |\Phi_\alpha(E)|\, d\alpha. \tag{1-7}$$

Our basic assumption on Γ is that it is a piecewise $\mathcal{C}^1$ curve with piecewise bi-Lipschitz continuous unit tangent vectors; these conditions will be discussed in the transversality analysis that appears in Section 2C, as well as in Section 4C, where we consider what goes wrong for nontransversal families.

The maps under consideration were originally introduced by Simon and the second author of this paper to study sum sets of the form $E+\Gamma$, where Γ denotes a sufficiently smooth curve and E denotes a compact set in $\mathbb{R}^2$. To see the connection, we write

$$\begin{aligned}\mathrm{Fav}_\Gamma(E) &= |\{(\alpha,\beta)\in\mathbb{R}^2 : \Phi_\alpha^{-1}(\beta)\cap E\neq\varnothing\}|\\ &= |\{(\alpha,\beta)\in\mathbb{R}^2 : \{(\alpha,\beta)-\Gamma\}\cap E\neq\varnothing\}|\\ &= |\{(\alpha,\beta)\in\mathbb{R}^2 : (\alpha,\beta)\cap(E+\Gamma)\neq\varnothing\}| = |E+\Gamma|.\end{aligned}$$

The measure and dimension of sets of the form $E+\Gamma$ was established in [Simon and Taylor 2022] and the interior of such sum sets was subsequently studied in [Simon and Taylor 2020]. Connections to the study of pinned distance sets and the Falconer distance conjecture are also explored there. In both [Simon and Taylor 2020; 2022], the results rely on relating the set E to the dimension, measure, and interior of the images of E under the maps $\{\Phi_\alpha\}$. A unifying ingredient in each of these works is the observation that the maps introduced in (1-6) are similar to orthogonal projection maps from the perspectives of measure, dimension, and interior.

As a further interpretation of the Favard curve length, there is a probabilistic interpretation. The Favard length of a set is comparable to its Buffon needle probability (that is, the probability that a long, thin needle dropped near the set intersects the set). In the nonlinear setting, the Favard curve length is comparable to the probability that a dropped curve meets the set — that is, the probability that $\Gamma\cap E\neq\varnothing$ after conditioning to the event that Γ lies near E. We denote this probability by $\mathcal{P}_\Gamma(E)$. In summary,

$$\mathrm{Fav}_\Gamma(E)\sim|E+\Gamma|\sim\mathcal{P}_\Gamma(E), \tag{1-8}$$

and our Theorem 1.5 gives a lower bound on these equivalent quantities.

Cladek, Davey, and Taylor [Cladek et al. 2022] obtained upper and lower bounds on the Favard curve length of $\mathcal{K}_n$, the n-th generation in the construction of the four-corner Cantor set:

$$\frac{1}{n}\lesssim\mathrm{Fav}_\Gamma(\mathcal{K}_n)\lesssim n^{-1/6+\delta}, \tag{1-9}$$

which by (1-8) implies upper and lower bounds on $|\mathcal{K}_n+\Gamma|\sim\mathcal{P}_\Gamma(\mathcal{K}_n)$. The lower bound relied on self-similarity and a square-counting argument adapted to the nonlinear setting. We will use energy methods to provide a simple alternative proof of the lower bound in (1-9) which holds in a more general setting and does not require self-similarity. See Corollary 1.9 for the details. Further, we obtain a higher-dimensional analogue of the lower bound in (1-9); this is the topic of the next section. We return to our discussion of Favard curve length in Section 2C after stating our main results.

It is worth remarking that other authors have studied related Buffon-type probability problems. In particular, Bond and Volberg [2011] considered lower bounds in the context of the intersection of $\mathcal{K}_n$ with large circles of radius n. In that context, the curves were adapted to the generation n, instead of having a fixed underlying curve.

1A3. *Favard surface length in* $\mathbb{R}^d$. The Favard curve length can also be formulated in a higher-dimensional setting, and we refer to the resulting quantity as the *Favard surface length*. Note that we still use the term "length" as we will consider a family of maps $\Phi_\alpha : \mathbb{R}^d \to \mathbb{R}$ and take the average of the 1-dimensional measures of the images of E under such maps. To the best of the authors' knowledge, this is the first article to define such a general notion of Favard length in higher dimensions.

Let $\Gamma = \Gamma_d$ denote a surface in $\mathbb{R}^d$. Given $\alpha \in \mathbb{R}^{d-1}$ and $\vec{x} =: (x_1, \ldots, x_d) \in \mathbb{R}^d$, let $\Phi_\alpha(\vec{x})$ denote the set of x_d-coordinates of the intersection of $\vec{x} + \Gamma$ with the line $\tilde{x} =: (x_1, \ldots, x_{d-1}) = \alpha$. That is,

$$\Phi_\alpha(\vec{x}) = \{\beta \in \mathbb{R} : (\alpha, \beta) \in (\vec{x} + \Gamma) \cap \{\tilde{x} = \alpha\}\}. \tag{1-10}$$

Given $\beta \in \mathbb{R}$, the inverse set $\Phi_\alpha^{-1}(\beta) = \{p \in \mathbb{R}^d : \beta \in \Phi_\alpha(p)\}$ is given by $(\alpha, \beta) - \mathcal{C}$. If Γ can be expressed as the graph of a function and $\Phi_\alpha(\vec{x}) \neq \varnothing$, then $\Phi_\alpha(\vec{x})$ is a singleton and we identify $\Phi_\alpha(\vec{x})$ with that point.

If $E \subset \mathbb{R}^d$, then the *Favard surface length* of E is defined by

$$\mathrm{Fav}_{\Gamma,d}(E) := |\{(\alpha, \beta) \in \mathbb{R}^d : \Phi_\alpha^{-1}(\beta) \cap E \neq \varnothing\}| = \int_{\mathbb{R}^{d-1}} |\Phi_\alpha(E)|\, d\alpha. \tag{1-11}$$

As was the case for the Favard curve length defined in the previous section, the Favard surface length of a set E is equivalent to the d-dimensional Lebesgue measure of the Minkowski sum:

$$\mathrm{Fav}_{\Gamma,d}(e) \sim |E + \Gamma|_d.$$

The quantity $\mathrm{Fav}_{\Gamma,d}(E)$ has a probabilistic interpretation in terms of a Buffon *surface* problem.

1B. ***Overview of transversality.*** It is known that nonlinear analogues of Besicovitch's and Marstrand's projection theorems hold for families of maps satisfying a *transversality condition*. A version of the Besicovitch projection theorem for transversal families can be found in [Hovila et al. 2012], and a quantitative version is developed in [Davey and Taylor 2022]. Marstrand's theorem is developed in the transversal setting in [Solomyak 1998, Theorem 5.1] and [Mattila 2015, Chapter 18]; see also Proposition 1.4.

The concept of transversality originated in [Pollicott and Simon 1995], where it was used to study the Hausdorff dimension of the attractors of a one-parameter family of iterated function systems. Solomyak [1995] then developed the transversality condition for the absolute continuity of invariant measures for a one-parameter family of iterated function systems. Moreover, in [Solomyak 1998] he combined the methods from [Pollicott and Simon 1995; Solomyak 1995] to establish a much more general transversality method for generalized projections. The next step was made by Peres and Schlag [2000], who further developed the method of transversality and gave a number of far-reaching applications. Such results have been utilized and further developed by a number of authors with extensive geometric applications. See, for instance, [Bourgain 2010; Cladek et al. 2022; Peres and Schlag 2000; Shmerkin 2020; Simon and Taylor 2020; 2022].

The transversality condition naturally arises when studying projection-type operators that do not overlap too much with each other, and this paper will explore the role transversality plays in developing energy estimates. The transversality condition addresses how, for distinct points x and y in the plane, the graphs $\{(\theta, \pi_\theta(x))\}$ and $\{(\theta, \pi_\theta(y))\}$ should behave at points of intersection. Roughly speaking, it says that if $\pi_\theta(x)$ and $\pi_\theta(y)$ are close for some value of θ, then they cannot remain close as θ changes. That is, the graphs cannot intersect tangentially, but must do so at a positive angle.

An alternative perspective on transversality will frequently come up in our techniques. If x and y are two fixed points, then the set of projections which cannot distinguish x and y must be rather small; placing this on the appropriate scale, this means that for each $\delta > 0$ there is an upper bound on the size of the set

$$\left\{\theta : \frac{|\pi_\theta(x) - \pi_\theta(y)|}{|x - y|} \le \delta\right\}.$$

Informally, this means that if π_θ is a randomly chosen projection then it will, with high probability, separate x and y on the projection side.

We now make precise our notion of transversality. The main objects are an indexed family of maps, a common domain and codomain equipped with measures, and a probability measure on the index set. In Section 2, we will place each of the families mentioned previously in the context of this definition and establish transversality with the appropriate parameters.

Definition 1.2 (nonlinear projections). For $1 \le m < n$, a *family of projection-type operators* will have the following objects associated to it:

- a domain Ω contained in $\mathbb{R}^n$,
- a codomain X contained in a Euclidean space, a nonnegative integer m, and a Borel measure h on X such that
$$h(B(x, \delta)) \gtrsim \delta^m$$
for all $x \in X$ and $\delta \in (0, 1)$,
- an indexing set A contained in an Euclidean space equipped with a compactly supported probability measure ψ,
- and a family of maps $\widetilde{\pi_\alpha} : \Omega \to X$ indexed by $\alpha \in A$ such that the function $(p, \alpha) \mapsto \widetilde{\pi_\alpha}(p)$ is continuous.

In order to be transversal, we will require that the family of projections satisfies a compatibility condition for different parameters:

Definition 1.3 (transversality). For a given $s \ge 0$, a family of maps $\{\widetilde{\pi_\alpha} : \alpha \in A\}$ satisfying Definition 1.2 is called *s-transversal* if there exist constants $c > 0$ and $\delta_0 > 0$ so that, for all distinct $x, y \in \Omega$ and $0 < \delta \le \delta_0$, we have

$$\psi\{\alpha : |\widetilde{\pi_\alpha}(x) - \widetilde{\pi_\alpha}(y)| \le \delta|x - y|\} < c \cdot \delta^m \cdot |x - y|^{m-s}, \tag{1-12}$$

or equivalently that

$$\psi\{\alpha : |\widetilde{\pi_\alpha}(x) - \widetilde{\pi_\alpha}(y)| \le \delta\} < c \cdot \frac{\delta^m}{|x - y|^s}. \tag{1-13}$$

Although this definition is written with a tunable parameter s, our most important case will be when the parameter s for transversality matches the dimension m of the target space; in this case, the transversality condition reduces to

$$\psi\{\alpha : |\widetilde{\pi_\alpha}(x) - \widetilde{\pi_\alpha}(y)| \le \delta|x - y|\} \lesssim \delta^m.$$

We note that our definition has some points in common with [Mattila 2015, Definition 18.1], but that we do not require smoothness of the projections nor derivative bounds of nonzero order.

1C. ***Main results.*** The uniting theme of our results is that for families of maps satisfying the transversality condition introduced in Definition 1.3, the energies associated to a measure μ will be closely related to the energies of the pushforward measures $\widetilde{\pi}_{\alpha\sharp}\mu$. As a demonstration of the techniques, we will begin by giving a brief formulation of part of the Marstrand projection theorem in the transversal setting: the dimension of a typical projection of a set with dimension $s < 1$ does not decrease. The proof of this fact, found in Section 3, demonstrates the utility of examining the energy of pushforward measures and is similar to the presentation in [Mattila 2015, Chapter 18]. (For the statement of the Marstrand projection theorem in the classic setting for orthogonal projections, as well as a formulation in higher dimensions, see [Mattila 2015, Section 5.3].)

Proposition 1.4 (nonlinear Marstrand theorem). *Suppose that $\{\widetilde{\pi}_\alpha : \alpha \in A\}$ is a family of maps into an m-dimensional space supporting a measure h, as in Definition 1.2. If E is a set with Hausdorff dimension $t \leq m$ and the family of projections is m-transversal, then for ψ-almost every $\alpha \in A$ we have*

$$\dim_{\mathcal{H}} \widetilde{\pi}_\alpha E = t. \tag{1-14}$$

Developing the energy techniques further, we give more general asymptotic lower bounds on the average size of a projection. The next theorem serves as a direct generalization of Mattila's result, Theorem 1.1.

Theorem 1.5 (average nonlinear projection length for neighborhoods). *With the notation of Definition 1.2, assume that $\{\widetilde{\pi}_\alpha : \alpha \in A\}$ is an m-transversal family of projections into an m-dimensional space. Fix a positive Borel probability measure μ supported on a compact set $F \subseteq \Omega$, so that*

$$\mu(B(x, r)) \lesssim r^t$$

for all $x \in \Omega$ and $0 < r < \infty$.

- *If $t < m$, then*
$$\int_A h(\widetilde{\pi}_\alpha F(r))\, d\psi(\alpha) \gtrsim r^{m-t}.$$
- *If $t = m$, then*
$$\int_A h(\widetilde{\pi}_\alpha F(r))\, d\psi(\alpha) \gtrsim (\log r^{-1})^{-1}.$$

As a first application, we can phrase Theorem 1.5 in the setting of radial projections and visibility defined in (1-4) and (1-5) respectively.

Theorem 1.6 (visibility for surfaces in $\mathbb{R}^n$). *Fix a set $E \subseteq \mathbb{R}^n$ of positive and finite s-dimensional Hausdorff measure, and consider a vantage set A which is a piecewise smooth $(n-1)$-dimensional surface equipped with Hausdorff measure; assume that for all $a \in A$ and $e \in E$ we have $|a - e| \lesssim 1$. Finally, assume that there exists a positive ρ such that for almost every $a \in A$ the tangent plane based at a does not pass within distance ρ of E. The following statements hold:*

- *The family of radial projections $\{P_a : a \in A\}$ is $(n-1)$-transversal.*
- *If $s < n - 1$, we have*
$$\int_A \operatorname{vis}(a, E(r))\, d\mathcal{H}^{n-1}(a) \gtrsim r^{n-1-s}.$$

- *If $s = n-1$, we have*

$$\int_A \operatorname{vis}(a, E(r))\, d\mathcal{H}^{n-1}(a) \gtrsim (\log r^{-1})^{-1}.$$

The first claim of Theorem 1.6 is established in Section 2B and the latter two claims are established in Section 4.

In a similar manner, we can put this result in the context of Favard curve length defined in (1-7). For curves in the plane, our techniques yield the following:

Theorem 1.7 (Favard curve length of neighborhoods). *Let E be a compact set in the plane and Γ a piecewise $\mathcal{C}^1$ curve with piecewise bi-Lipschitz continuous unit tangent vectors. Assume further that E supports a Borel probability measure μ with the t-dimensional growth condition $\mu(B(x, r)) \lesssim r^t$ for all $x \in E$, $0 < r < \infty$. The following statements hold*:

- *The family of curve projections Φ_α is 1-transversal.*
- *If $t < 1$, then for all sufficiently small r we have*

$$\operatorname{Fav}_\Gamma(E(r)) \gtrsim r^{1-t}.$$

- *If $t = 1$, then for all sufficiently small r we have*

$$\operatorname{Fav}_\Gamma(E(r)) \gtrsim (\log r^{-1})^{-1}.$$

Next, we consider applications of Theorem 1.5 to study self-similar sets such as $\mathcal{K}_n$, the n-th generation in the construction of the four-corner Cantor set. Although they are not precisely the same as neighborhoods of 1-sets, the sets $\mathcal{K}_n$ still support measures with easily computable density and Mattila's energy techniques can be adapted to estimate their visibilities (1-5) and Favard curve lengths (1-7) from below. Our techniques are similarly amenable to such sets, and we will have the following corollaries:

Corollary 1.8 (visibility of $\mathcal{K}_n$). *Suppose that Γ is a smooth curve such that for any point $x \in [0, 1]^2$ and any $\gamma \in \Gamma$ we have $|x - \gamma| \sim 1$, and that no tangent line to Γ passes through $[0, 1]^2$. Then*

$$\int_\Gamma \operatorname{vis}(a, \mathcal{K}_n)\, d\mathcal{H}^1(a) \gtrsim \frac{1}{n}.$$

Corollary 1.9 (Favard curve length of $\mathcal{K}_n$). *If Γ is a piecewise $\mathcal{C}^1$ curve with piecewise bi-Lipschitz continuous unit tangent vectors, then*

$$\operatorname{Fav}_\Gamma(\mathcal{K}_n) \gtrsim \frac{1}{n}.$$

Although these results are stated for the generations $\mathcal{K}_n$ specifically, there are substantial generalizations of the results. The core fact used in the proof is that $\mathcal{K}_n$ supports a measure with a specific density property; this behavior can be observed in a very broad family of 1-dimensional fractal sets generated by iterated function systems.

Finally, we consider an application of Theorem 1.5 for the Favard surface length, defined in (1-11), when $d = 3$. Although we do not state them here, there are natural generalizations of this result to arbitrary dimension.

Theorem 1.10 (Favard surface length of neighborhoods). *Let E be a compact set in the plane and Γ denote a surface in $\mathbb{R}^3$ defined by $\Gamma = \{(t, \gamma(t)) : t \in I\}$, where $\gamma : \mathbb{R}^2 \to \mathbb{R}$, $\gamma(s) = f(|s|)$, and $f : \mathbb{R} \to \mathbb{R}$ is a C^2 function on a nonempty compact interval I satisfying $f(x) = f(-x)$, with $f'' > 0$ on I. Assume further that E supports a Borel probability measure μ with the t-dimensional growth condition $\mu(B(x, r)) \lesssim r^t$ for all $x \in E$, $0 < r < \infty$. The following statements hold*:

- *The family of curve projections Φ_α is 1-transversal.*
- *If $t < 1$, then for all sufficiently small r we have*

$$\mathrm{Fav}_\Gamma(E(r)) \gtrsim r^{1-t}.$$

- *If $t = 1$, then for all sufficiently small r we have*

$$\mathrm{Fav}_\Gamma(E(r)) \gtrsim (\log r^{-1})^{-1}.$$

The outline of the paper is as follows. In Section 2, we will show how each of the aforementioned families of maps exhibit the required transversality properties. Geometrically motivated proofs are given for each family. Section 3 develops the energy techniques necessary to study pushforward measures, beginning with an illustration of how a transversal family of maps can be used to prove a classical result of Marstrand. The proof of Theorem 1.5 appears in Section 3. In Section 4, we prove Theorems 1.6 and 1.7 as applications of Theorem 1.5 paired with the transversality established in Section 2, and we explore applications and sharpness examples.

2. Establishing transversality

The aim of this section is to illustrate several families of projections that meet the transversality condition described in Definition 1.3. This includes orthogonal, radial, curve, and surface projections.

2A. ***Orthogonal projections.*** Our first example of a transversal family is the collection of orthogonal projections from $\mathbb{R}^n$ to $\mathbb{R}^m$ for some $m < n$. To be explicit about the setup, we will consider a domain $\Omega = \mathbb{R}^n$, a codomain $X = \mathbb{R}^m$, and equip the codomain with the appropriate Lebesgue measure. We then have the family

$$\{\iota_V \circ P_V : V \in G(n, m)\}$$

of projections indexed by the Grassmanian, where P_V is the orthogonal projection into the m-plane V, and with the natural inclusion $\iota_V : V \to \mathbb{R}^m$; equip this set with the Haar measure $\gamma_{n,m}$. The full details of the construction of the Grassmanian manifold and the measure $\gamma_{n,m}$ can be found, for example, in [Mattila 1995, Chapter 3].

To establish transversality, the core estimate in this context is contained in [Mattila 1995, Lemma 2.7]: for any distinct points $x, y \in \mathbb{R}^n$,

$$\gamma_{n,m}(\{V \in G(n, m) : |P_V(x - y)| \le \delta\}) \sim \frac{\delta^m}{|x - y|^m}. \tag{2-1}$$

Using the linearity of P_V, one can quickly establish:

Lemma 2.1 (orthogonal projections are transversal). *The family of orthogonal projections from $\mathbb{R}^n$ to $\mathbb{R}^m$ equipped with the Haar measure $\gamma_{n,m}$ is m-transversal.*

As in [Mattila 1995], this can be done geometrically, by reducing to an estimate of the m-dimensional measure of a patch on a sphere. There is also an important probabilistic interpretation, which will turn out to be the main ingredient when studying other transversal families. If x and y are fixed points in $\mathbb{R}^n$, then a randomly chosen m-dimensional plane is likely to preserve some, if not most, of the distance between x and y; that is, on average we have that $|P_V(x) - P_V(y)| \geq \delta|x-y|$. However, there is still an exceptional set of m-planes which do not respect this inequality at scale δ — for example, any m-plane which is sufficiently close to lying in the orthogonal complement to the line between x and y. Transversality comes from controlling the $\gamma_{n,m}$-measure of the exceptional set for scale δ.

2B. ***Visibility.*** We now turn to establishing the transversality condition for families of radial maps. We begin by first recalling the notation defined in Section 1A1. For a point a in $\mathbb{R}^n$, the radial projection based at a maps $\mathbb{R}^n \setminus \{a\} \to \mathbb{S}^{n-1}$ via

$$P_a(x) = \frac{x-a}{|x-a|}.$$

For a fixed vantage set $A \subset \mathbb{R}^n$ equipped with a measure ψ, our family of projections will be $\{P_a : a \in A\}$. The common domain will be a visible set E, which will be assumed to be disjoint from A. Our codomain is $\mathbb{S}^{n-1}$ equipped with the surface measure and so $m = n-1$ and $P_a : E \to S^{n-1}$. The aim of this section is to establish some minimal geometric relations between the vantage set A with the measure ψ and the visible set E so that the family $\{P_a : a \in A\}$ is $(n-1)$-transversal. A natural condition on the probability measure ψ will arise after we analyze the geometry of the radial projections.

To this end, we will make use of the following geometric lemma. A 2-dimensional variant appeared in [Bond et al. 2016, Lemma 2.3]; we will provide a somewhat different proof and generalize the result to higher dimensions.

Lemma 2.2 (visibility and tubes). *Fix a scale $R > 0$ and two points x, y not contained in the vantage set A with $|x-y| \leq R$. Let $L_{x,y}$ denote the line connecting them. Then there exists a constant $C < \infty$ depending only on R such that*

$$\{a \in A : |P_a(x) - P_a(y)| \leq \delta|x-y|\} \cap B(x, R) \subseteq L_{x,y}(C\delta),$$

where $L_{x,y}(C\delta)$ denotes the $C\delta$-neighborhood of the line $L_{x,y}$.

Proof. We proceed by contrapositive. Suppose that a is within the ball $B(x, R)$ but outside the tube $L_{x,y}(\rho)$ of radius ρ around $L_{x,y}$. Draw a triangle with vertices x, y, and a; let θ denote the internal angle at vertex a and γ denote the internal angle at vertex y. Since $|P_a(x) - P_a(y)|$ is comparable to the internal angle θ of the triangle, it is sufficient to give a lower bound on the angle θ. By the law of sines, we have

$$\frac{\sin\theta}{|x-y|} = \frac{\sin\gamma}{|a-x|}$$

so that

$$\theta \geq \sin\theta = \frac{\sin\gamma}{|a-x|}|x-y|.$$

If $\mathcal{A}_{x,y,a}$ denotes the altitude of the triangle (as viewed with base side $\overline{xy}$) then

$$\sin\gamma = \frac{\mathcal{A}_{x,y,a}}{|y-a|}$$

and

$$\theta \geq \frac{\mathcal{A}_{x,y,a}\cdot|x-y|}{|a-x|\cdot|a-y|}.$$

Since $a, x, y \in B(x, R)$, we have $|a-x| \leq 2R$ and $|a-y| \leq 2R$. Since a lies outside the tube $L_{x,y}(\rho)$, the altitude must be at least ρ. Therefore, there exists a constant $c \sim 1$ for which

$$|P_a(x) - P_a(y)| \geq c\theta \geq c\cdot\frac{\rho}{4R^2}\cdot|x-y|.$$

Choosing $\rho = C\delta$ for $C > 4R^2/c$ establishes that $|P_a(x) - P_a(y)| > \delta|x-y|$, as desired. □

We now have a natural condition to impose on the probability measure ψ: as we wish to verify (1-12) with $s = m = n-1$, Lemma 2.2 implies that the measure of a tube should be bounded by the radius of the tube to an appropriate power. To be precise, we will say that ψ satisfies the *tube condition with respect to E* if for any tube T_δ with sufficiently small radius δ that passes through the visible set, E, we have

$$\psi(T_\delta) \lesssim \delta^{n-1}. \tag{2-2}$$

In this case, provided the distance from A to E is at most R, we have established that $\{P_a : a \in A\}$ is a family of maps from an n-dimensional space to an $(n-1)$-dimensional space with

$$\psi\{a \in A : |P_a(x) - P_a(y)| \leq \delta|x-y|\} \lesssim \delta^{n-1}.$$

Comparing this to the definition of transversality, we have established the following:

Lemma 2.3 (radial maps are transversal). *Fix a scale $R > 0$. Fix a vantage set A and a visible set E with the condition that for all $a \in A$ and $e \in E$ we have $|a-e| \lesssim 1$. If A is equipped with a measure ψ satisfying the tube condition with respect to E (2-2), then the family $\{P_a : a \in A\}$ is $(n-1)$-transversal as in (1-13).*

This gives a substantial degree of flexibility in structuring the vantage set. One application of this technique is to a vantage set which is made up of a smooth curve Γ whose tangent lines do not come too close to the visible set. When ψ is taken to be the restriction of $\mathcal{H}^{n-1}$ to the vantage set A, this will imply that ψ satisfies the tube condition with respect to E. We discuss this idea more in Section 4A.

2C. *Favard curve length.* In this section, we verify that the maps $\Phi_\lambda : \mathbb{R}^2 \to \mathbb{R}$ introduced in (1-6) satisfy the transversality condition of Definition 1.3. This will proceed through a couple of reductions. First, we will set up some basic assumptions on the smoothness of the curve, as well as some notation. Next, by breaking the curve into simpler pieces, we reduce to the case of a curve that is a graph satisfying a simpler curvature condition. We establish transversality in this simpler setting and note this is sufficient to establish lower bounds on the Favard length for the general setting.

Definition 2.4. We say that Γ satisfies our *standard curvature condition* if Γ is a piecewise $\mathcal{C}^1$ curve with piecewise bi-Lipschitz continuous unit tangent vectors.

Under the assumptions of Definition 2.4, Γ can be expressed as a disjoint union of continuous subcurves $\Gamma = \bigcup_{i=1}^{\infty} \Gamma_i$, where each Γ_i is $\mathcal{C}^1$ of finite length with a bi-Lipschitz continuous unit tangent vector. By further decomposition of the curve, each Γ_i can be expressed either as a graph with respect to the first coordinate, $\Gamma_i = \{(t, \gamma_i(t)) : t \in I_i\}$, or as a graph with respect to the second coordinate, $\Gamma_i = \{(\gamma_i(t), t) : t \in I_i\}$, so that $\sup_{t \in I_i} |\gamma_i'(t)| \le 1$, and γ_i' is λ_i-bi-Lipschitz.

In order to obtain lower bounds on $\mathrm{Fav}_\Gamma(E)$, where E will denote a compact subset of $\mathbb{R}^2$, since $\mathrm{Fav}_\Gamma(E) \ge \mathrm{Fav}_{\Gamma_i}(E)$ for each i, it suffices to obtain lower bounds on $\mathrm{Fav}_{\Gamma_i}(E)$. Fixing i and observing that rotating the curve and the set E by the same amount has no affect on $\mathrm{Fav}_{\Gamma_i}(E) = |E + \Gamma_i|$, we may simply assume that Γ_i is a graph with respect to the first coordinate. Finally, for ease of notation, we drop the subscript i and assume that Γ has all the properties of Γ_i.

Definition 2.5. We say that Γ is a curve satisfying the *simple curvature condition* if $\Gamma = \{(t, \gamma(t)) : t \in I\}$, where $\gamma : \mathbb{R} \to \mathbb{R}$,

$$\sup_{t \in I} |\gamma'(t)| \le 1, \tag{2-3}$$

and γ' is Λ-bi-Lipschitz satisfying

$$\Lambda^{-1}|s - t| \le |\gamma'(s) - \gamma'(t)| \le \Lambda |s - t| \tag{2-4}$$

for some $0 < \Lambda < \infty$ and for each s, t in a nontrivial closed interval I.

Let $\Gamma = \{(t, \gamma(t)) : t \in I\}$ be a curve satisfying the simple curvature condition of Definition 2.5. Note that condition (2-4) guarantees that γ' is monotonic; without loss of generality, we will assume that Γ is concave down so that if $t < s$, then

$$\frac{\gamma'(s) - \gamma'(t)}{s - t} < 0. \tag{2-5}$$

Write $I = [L_1, L_2]$ for some $L_1 < L_2$ and set $h = \frac{1}{2}(L_2 - L_1)$. Set $\Omega = [0, h]^2 \subset \mathbb{R}^2$ and $A = [L_1 + h, L_2]$. With this set up, for each $\lambda \in A$ and $\boldsymbol{a} \in \Omega$,

$$\ell_\lambda \cap (\boldsymbol{a} + \Gamma) = (\lambda, a_2 + \gamma(\lambda - a_1))$$

is a singleton, as in Figure 1, and we can define the one-parameter family of mappings $\{\Phi_\lambda(\cdot)\}_{\lambda \in A}$, $\Phi_\lambda : \Omega \to \ell_\lambda$, by

$$\Phi_\lambda(\boldsymbol{a}) = a_2 + \gamma(\lambda - a_1). \tag{2-6}$$

We are now ready to show that the simple curvature assumption implies 1-transversality. In line with Definition 1.2, our codomain is $\mathbb{R}$ equipped with the 1-dimensional Lebesgue measure and so $m = 1$.

Lemma 2.6 (curve maps are transversal). *Let Γ be a curve satisfying the simple curvature assumption of Definition 2.5. Equip the parameter space A with the 1-dimensional Lebesgue measure. Then the associated family of projections $\{\Phi_\lambda : \Omega \to \mathbb{R} : \lambda \in A\}$ is 1-transversal as in (1-13).*

Proof. Fix a choice of $\boldsymbol{a} = (a_1, a_2)$, $\boldsymbol{b} = (b_1, b_2) \in \Omega$, with $\boldsymbol{a} \ne \boldsymbol{b}$. The proof comes in two parts: the translated graphs $(\boldsymbol{a} + \Gamma)$ and $(\boldsymbol{b} + \Gamma)$ will either intersect at a point, or they will be disjoint. We first

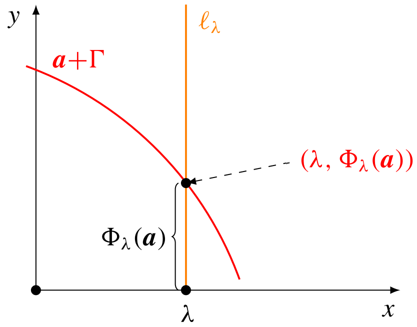

Figure 1. $\Phi_\lambda(\boldsymbol{a})$.

handle the intersecting case when

$$(\boldsymbol{a}+\Gamma)\cap(\boldsymbol{b}+\Gamma)\neq\varnothing. \tag{2-7}$$

That is, suppose there exist $s_0, t_0 \in I$ and $\boldsymbol{a} = (a_1, a_2) \in \mathbb{R}^2$ such that

$$\boldsymbol{x} := (a_1, a_2) + (s_0, \gamma(s_0)) = (b_1, b_2) + (t_0, \gamma(t_0)).$$

Comparing coordinates, we have

$$x_1 = a_1 + s_0 = b_1 + t_0, \tag{2-8}$$

$$x_2 = a_2 + \gamma(s_0) = b_2 + \gamma(t_0). \tag{2-9}$$

For $\lambda \in A$, set

$$d_\lambda := \operatorname{dist}(\boldsymbol{x}, \ell_\lambda) = |\lambda - x_1|, \tag{2-10}$$

as depicted in Figure 2. We verify that

$$|\Phi_\lambda(\boldsymbol{a}) - \Phi_\lambda(\boldsymbol{b})| \sim d_\lambda \cdot |\boldsymbol{a}-\boldsymbol{b}|, \tag{2-11}$$

where the implied constant is independent of λ, $\boldsymbol{a}$, and $\boldsymbol{b}$. Strictly speaking, we only need that the left-hand side dominates the right-hand side. Upon establishing (2-11), it will follow that if $\delta > 0$ and $\lambda \in A$ satisfy $|\Phi_\lambda(\boldsymbol{a}) - \Phi_\lambda(\boldsymbol{b})| \leq \delta$, then

$$d_\lambda \cdot |\boldsymbol{a}-\boldsymbol{b}| \lesssim \delta,$$

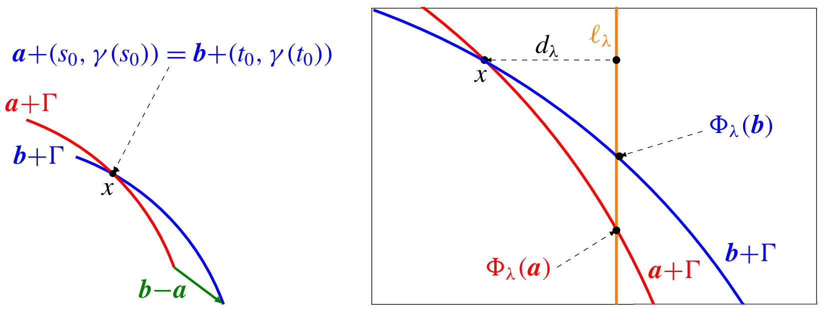

Figure 2

and so

$$|\{\lambda \in A : |\Phi_\lambda(\boldsymbol{a}) - \Phi_\lambda(\boldsymbol{b})| \le \delta\}| \lesssim \frac{\delta}{|\boldsymbol{a}-\boldsymbol{b}|} \tag{2-12}$$

which is the desired transversality condition.

We have two further reductions. First, as depicted in Figure 2, we consider the case when $\lambda \ge x_1$ so that

$$d_\lambda = \lambda - x_1 \ge 0. \tag{2-13}$$

Note that the case when $\lambda - x_1 < 0$ can be handled by reflecting E and Γ about the y-axis. Secondly, by relabeling $\boldsymbol{a}$ and $\boldsymbol{b}$ if necessary, we may assume that when $\lambda > x_1$, we have

$$\Phi_\lambda(\boldsymbol{b}) - \Phi_\lambda(\boldsymbol{a}) > 0 \tag{2-14}$$

as in Figure 2. Finally, we will also have

$$(b_1 - a_1) > 0. \tag{2-15}$$

This follows from the geometry of the curves: in order for (2-14) to hold in the intersecting case, the convexity of Γ shows that $\boldsymbol{b}$ must lie below and to the right of $\boldsymbol{a}$.

Using the convexity condition (2-5), we will show that

$$\Phi_\lambda(\boldsymbol{b}) - \Phi_\lambda(\boldsymbol{a}) \sim (b_1 - a_1)\cdot d_\lambda. \tag{2-16}$$

Observe that by the bound on γ' and the relationships established in (2-8) and (2-9),

$$|b_2 - a_2| = |\gamma(s_0) - \gamma(t_0)| \le |s_0 - t_0| = |b_1 - a_1|. \tag{2-17}$$

As such, proving (2-16) will be sufficient to establish (2-11). We now carry out the verification of (2-16) in three cases based on the relative sizes of d_λ and $|b_1 - a_1|$. We will handle the nonintersecting case (where (2-7) does not hold) separately.

<u>Case 1</u>: $(b_1 - a_1) < \frac{1}{2} d_\lambda$. We begin by examining the simplest case, which motivates the finer analysis to come. This is depicted in Figure 3.

Using the relationships established in (2-6)–(2-9) and the mean value theorem, we have

$$\begin{aligned}
\Phi_\lambda(\boldsymbol{b}) - \Phi_\lambda(\boldsymbol{a}) &= (b_2 + \gamma(\lambda - b_1)) - (a_2 + \gamma(\lambda - a_1)) \\
&= (b_2 - a_2) + (\gamma(\lambda - b_1) - \gamma(\lambda - a_1)) \\
&= (\gamma(s_0) - \gamma(t_0)) + (\gamma(\lambda - b_1) - \gamma(\lambda - a_1)) \\
&= \gamma'(\xi)(b_1 - a_1) - \gamma'(\eta)(b_1 - a_1) \\
&= [\gamma'(\xi) - \gamma'(\eta)](b_1 - a_1)
\end{aligned}$$

for some $\eta \in (\lambda - b_1, \lambda - a_1)$ and $\xi \in (t_0, s_0)$. It follows by (2-4) and (2-5) that

$$\Phi_\lambda(\boldsymbol{b}) - \Phi_\lambda(\boldsymbol{a}) \sim (\eta - \xi)\cdot(b_1 - a_1).$$

Since $(b_1 - a_1) < \frac{1}{2} d_\lambda$, we see that (2-16) is verified following the observation that

$$(\eta - \xi) \sim d_\lambda.$$

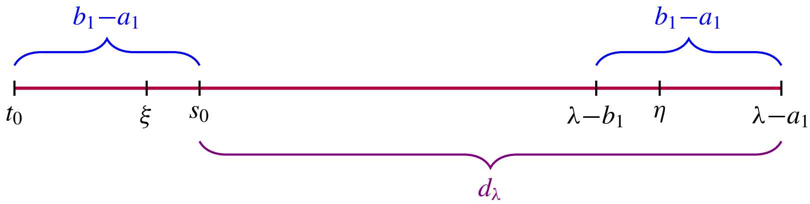

Figure 3. Case 1.

To see this, recall from (2-13) and (2-8) that $d_\lambda = \lambda - a_1 - s_0 = \lambda - b_1 - t_0$. Following Figure 3,

$$\eta - \xi > (\lambda - b_1) - s_0 = (\lambda - b_1 - t_0) - (s_0 - t_0) = d_\lambda - (b_1 - a_1),$$

and similarly

$$\eta - \xi < (\lambda - a_1) - t_0 = (\lambda - a_1 - s_0) + (s_0 - t_0) = d_\lambda + (b_1 - a_1).$$

Before moving to the general argument, we observe that the separation of d_λ and $(b_1 - a_1)$ was crucial in guaranteeing that the variables arising from the application of the mean value theorem, ξ and η, were properly separated. More generally, a finer analysis using telescoping sums is used to guarantee such separation.

Case 2: $\frac{1}{2} d_\lambda \leq (b_1 - a_1) < d_\lambda$. Set

$$p = \tfrac{1}{2}(b_1 - a_1) \quad \text{and} \quad q = s_0. \tag{2-18}$$

First, we take a moment to compare the variables under examination. Note $p > 0$ by (2-15). Using (2-13) and (2-8), we can write $d_\lambda = \lambda - b_1 - t_0$ and $b_1 - a_1 = s_0 - t_0$. Therefore, when $b_1 - a_1 < d_\lambda$, we have $s_0 - t_0 < \lambda - b_1 - t_0$ and so $s_0 < \lambda - b_1$. This implies

$$t_0 < s_0 < \lambda - b_1 < \lambda - a_1,$$

and so, for p and q as in (2-18),

$$t_0 = q - 2p < q - p < q = s_0 < \lambda - b_1 = \lambda - a_1 - 2p < \lambda - a_1 - p < \lambda - a_1.$$

Appealing to (2-6) and (2-9), we can write

$$\begin{aligned}
\Phi_\lambda(\boldsymbol{a}) - \Phi_\lambda(\boldsymbol{b}) &= \gamma(\lambda - a_1) - \gamma(\lambda - b_1) - (b_2 - a_2) \\
&= \gamma(\lambda - a_1) - \gamma(\lambda - b_1) - (\gamma(s_0) - \gamma(t_0)) \\
&= \sum_{j=0}^{1} \big(\gamma(\lambda - a_1 - jp) - \gamma(\lambda - a_1 - (j+1)p)\big) - \sum_{j=0}^{1} \big(\gamma(q - jp) - \gamma(q - (j+1)p)\big).
\end{aligned}$$

Applying the mean value theorem, there exist $h_0, h_1, h'_0, h'_1 \in (0, 1)$ so that

$$\Phi_\lambda(\boldsymbol{a}) - \Phi_\lambda(\boldsymbol{b}) = \sum_{j=0}^{1} (\gamma'(\lambda - a_1 - jp - h_j p) \cdot p) - \sum_{j=0}^{1} (\gamma'(q - jp - h'_j p) \cdot p),$$

and it follows that

$$\Phi_\lambda(\boldsymbol{a}) - \Phi_\lambda(\boldsymbol{b}) \sim \left(\sum_{j=0}^{1} (\gamma'(\lambda - a_1 - jp - h_j p) - \gamma'(q - jp - h'_j p)) \right) \cdot p. \tag{2-19}$$

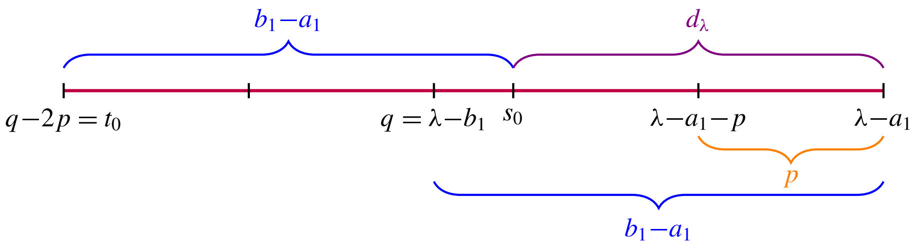

Figure 4. Case 3.

The purpose for adding and subtracting terms, is that the terms $\lambda-a_1-jp-h_jp$ and $q-jp-h'_jp$ are now appropriately separated for $j=0,1$. Indeed, when $d_\lambda>(b_1-a_1)$, recalling that $q=s_0$, it holds that

$$(\lambda-a_1-jp-h_jp)-(s_0-jp-h'_jp)=d_\lambda-h_jp+h'_jp\geq\tfrac{1}{2}d_\lambda,$$

and

$$(\lambda-a_1-jp-h_jp)-(s_0-jp-h'_jp)=d_\lambda-h_jp+h'_jp\leq\tfrac{3}{2}d_\lambda.$$

The key point is that in (2-19), the arguments of γ' within each summand are separated by a positive quantity comparable to d_λ. Using the bi-Lipschitz condition on γ' (in which case γ' is strictly monotonic on I), we conclude that

$$\Phi_\lambda(\boldsymbol{b})-\Phi_\lambda(\boldsymbol{a})\sim d_\lambda\cdot p.$$

Since $p\sim(b_1-a_1)$, this case is completed.

Case 3: $d_\lambda\leq(b_1-a_1)$. Set

$$p=\tfrac{1}{2}d_\lambda\quad\text{and}\quad q=(\lambda-b_1).\tag{2-20}$$

With this choice of p and q, the proof proceeds as in the previous case. This situation is depicted in Figure 4.

Using (2-8) and (2-13), we can write $d_\lambda=\lambda-b_1-t_0=\lambda-a_1-s_0\geq 0$ and $b_1-a_1=s_0-t_0>0$. Therefore, when $d_\lambda\leq b_1-a_1$, we have $\lambda-b_1-t_0\leq s_0-t_0$ and so $\lambda-b_1\leq s_0$. Combining these observations, if $d_\lambda\leq(b_1-a_1)$, then

$$t_0\leq\lambda-b_1\leq s_0\leq\lambda-a_1,$$

and so, for p and q as in (2-20),

$$t_0=q-2p\leq q-p\leq q=\lambda-b_1\leq s_0=\lambda-a_1-2p\leq\lambda-a_1-p\leq\lambda-a_1.$$

Using an identical telescoping argument to that used in the previous case to obtain (2-19), except now with p and q as in (2-20), we conclude that there exist $h_0,h_1,h'_0,h'_1\in(0,1)$ so that

$$\Phi_\lambda(\boldsymbol{a})-\Phi_\lambda(\boldsymbol{b})\sim\left(\sum_{j=0}^{1}(\gamma'(\lambda-a_1-jp-h_jp)-\gamma'(q-jp-h'_jp))\right)\cdot p.\tag{2-21}$$

We now observe that $\lambda - a_1 - jp - h_j p$ and $q - jp - h'_j p$ are sufficiently separated for $j = 0, 1$ when $d_\lambda \leq b_1 - a_1$:

$$(\lambda - a_1 - jp - h_j p) - (q - jp - h'_j p) = b_1 - a_1 - h_j p + h'_j p \geq \tfrac{1}{2}(b_1 - a_1)$$

and

$$(\lambda - a_1 - jp - h_j p) - (q - jp - h'_j p) = b_1 - a_1 - h_j p + h'_j p \leq \tfrac{3}{2}(b_1 - a_1).$$

As in Case 2 above, we have now established the necessary separation between the arguments of γ' in each summand; it follows that

$$\Phi_\lambda(\boldsymbol{b}) - \Phi_\lambda(\boldsymbol{b}) \sim (b_1 - a_1) \cdot p.$$

Since $p \sim d_\lambda$, this case is finished.

Nonintersection case: It remains to verify (1-13) when (2-7) does not hold. Assume that $\boldsymbol{a}$ and $\boldsymbol{b}$ are such that

$$(\boldsymbol{a} + \Gamma) \cap (\boldsymbol{b} + \Gamma) = \varnothing. \tag{2-22}$$

Let $\delta > 0$. For each $\lambda \in A$, set

$$h(\lambda) := \Phi_\lambda(\boldsymbol{b}) - \Phi_\lambda(\boldsymbol{a}) = \gamma(\lambda - b_1) - \gamma(\lambda - a_1) + (b_2 - a_2).$$

Relabeling if necessary, we may assume that the graph $(\boldsymbol{b} + \Gamma)$ is above $(\boldsymbol{a} + \Gamma)$ in the sense that, for each $\lambda \in A$, it holds that

$$h(\lambda) > 0.$$

Observe that in the case $a_1 = b_1$, we have $h(\lambda) = b_2 - a_2$ is constant, and so the left-hand-side of (1-13) is nonzero identically when $|a - b| = |a_2 - b_2| \leq \delta$, in which case the right-hand-side of (1-13) is bounded below by the constant c, and the inequality is satisfied provided that c is chosen so that $c \geq |A|$.

Assume then that $a_1 \neq b_1$. We will apply a vertical shift to the curve $(\boldsymbol{b} + \Gamma)$ to reduce to the intersection case considered in (2-7) and handled above. It is a consequence of the curvature assumption of Definition 2.5 that there exists a unique $\hat{\lambda} \in A$ where $h(\lambda)$ is minimized. Set

$$d := h(\hat{\lambda}).$$

(Indeed, when $a_1 \neq b_1$, note that h is strictly monotonic as $h' \neq 0$ by (2-4)). Now

$$(\Gamma + (b_1, b_2 - d)) \cap (\Gamma + \boldsymbol{a}) \neq \varnothing,$$

and we see that

$$\Phi_\lambda(\boldsymbol{b}) = b_2 + \gamma(\lambda - b_1) = b_2 - d + \gamma(\lambda - b_1) + d = \Phi_\lambda((b_1, b_2 - d)) + d. \tag{2-23}$$

Set $\boldsymbol{b}(d) = (b_1, b_2 - d)$. Now, if λ is such that $h(\lambda) = \Phi_\lambda(\boldsymbol{b}) - \Phi_\lambda(\boldsymbol{a}) \leq \delta$, then

$$\Phi_\lambda(\boldsymbol{b}(d)) - \Phi_\lambda(\boldsymbol{a}) \leq \delta - d \leq \delta.$$

Note we may assume that $\delta \geq d$ since $h(\lambda) \geq d$ for each $\lambda \in A$. Therefore

$$\{\lambda \in A : \Phi_\lambda(\boldsymbol{b}) - \Phi_\lambda(\boldsymbol{a}) \leq \delta\} \subset \{\lambda \in A : \Phi_\lambda((b_1, b_2 - d)) - \Phi_\lambda(\boldsymbol{a}) \leq \delta\}, \tag{2-24}$$

and it follows from the previous Cases 1–3 that there exists a constant $c > 0$ that depends only on the constant Λ in (2-4) so that

$$|\{\lambda \in A : \Phi_\lambda((b_1, b_2 - d)) - \Phi_\lambda(\boldsymbol{a}) \le \delta\}| \le \frac{c\,\delta}{|\boldsymbol{b}(d) - \boldsymbol{a}|}. \tag{2-25}$$

Combining (2-24) and (2-25), we see that if $|\boldsymbol{b}(d)-\boldsymbol{a}|$ were bounded below by $|\boldsymbol{b}-\boldsymbol{a}|$, then the argument would be complete. Since this may not always be the case, we need a slightly more delicate analysis.

We will now proceed in two cases, based on the relative sizes of $|b_1 - a_1|$ and $|b_2 - a_2|$. When the first difference is dominant, the shift between $\boldsymbol{b}$ and $\boldsymbol{a}$ is mostly horizontal and this horizontal translation is detected by the first coordinate of $\boldsymbol{b}(d)$. The more challenging case is when the translation is nearly vertical; this will follow the same lines as when $b_1 = a_1$. To be precise, we now consider the cases when $|b_1 - a_1| \ge \frac{1}{2}|b_2 - a_2|$ and $|b_1 - a_1| < \frac{1}{2}|b_2 - a_2|$ separately.

In the former case,

$$|b_1 - a_1| \gtrsim |b - a|$$

and so

$$|\boldsymbol{b}(d) - \boldsymbol{a}|^2 = |b_1 - a_1|^2 + |b_2 - d - a_2|^2 \ge |b_1 - a_1|^2 \gtrsim |b - a|^2.$$

In this case, we see that $|\boldsymbol{b}(d) - \boldsymbol{a}|$ is bounded below by a constant multiple of $|\boldsymbol{b} - \boldsymbol{a}|$, and the argument is complete upon combining (2-24) and (2-25).

Now consider the latter case that $|b_1 - a_1| < \frac{1}{2}|b_2 - a_2|$. Suppose that λ is such that $h(\lambda) \le \delta$. By the mean value theorem, there exists an η so that

$$\gamma(\lambda - b_1) - \gamma(\lambda - a_1) = -\gamma'(\eta)(b_1 - a_1).$$

Recall from (2-3) that $\sup_{t\in I} |\gamma'(t)| \le 1$. It follows from the reverse triangle inequality that

$$\begin{aligned} h(\lambda) &\ge |b_2 - a_2| - |\gamma'(\eta)(b_1 - a_1)| \\ &\ge |b_2 - a_2| - |b_1 - a_1| \\ &\ge |b_2 - a_2| - \tfrac{1}{2}|b_2 - a_2| \\ &= \tfrac{1}{2}|b_2 - a_2| \sim |\boldsymbol{b} - \boldsymbol{a}|, \end{aligned}$$

where the implicit constants are independent of $\boldsymbol{b}$, $\boldsymbol{a}$ and λ. It follows that there exists a $c' > 0$ so that if λ is such that $h(\lambda) \le \delta$, then $|\boldsymbol{b} - \boldsymbol{a}| \le c'\delta$ or $1 \le c'\delta/|\boldsymbol{b} - \boldsymbol{a}|$. Now,

$$|\{\lambda \in A : h(\lambda) \le \delta\}| \le |A| \le c \le c\frac{c'\,\delta}{|\boldsymbol{b} - \boldsymbol{a}|},$$

provided c is chosen so that $c \ge |A|$. □

2D. *Surface projections.* Here, we show that the maps corresponding to the Favard surface length and introduced in Section 1A3 satisfy the transversality condition of (1-13). We will consider the case when Γ is a surface of revolution generated by an even, C^2, concave-up function f defined on a neighborhood of the origin. That is, Γ will be the graph of $\gamma : \mathbb{R}^2 \to \mathbb{R}$ given by

$$\gamma(s) = f(|s|)$$

defined on a closed ball $B := \overline{B(0, L)}$ for some $L > 0$.

Note that $f'' > 0$ on $[-L, L]$. It is straightforward to check that $f'(x) \geq 0$ on $[0, L]$ with equality only at $x = 0$; computing the second partial derivative of γ in x at $|(x, y)| = 0$ and $|(x, y)| \neq 0$ separately shows that there exists $c > 0$ so that for each $(x, y) \in B$

$$\frac{\partial^2 \gamma}{\partial x^2}(x, y) > c. \tag{2-26}$$

Now, we choose a parameter set A and a domain Ω as in Definition 1.2: set $A = \overline{B\big(0, \frac{1}{3}L\big)} \subset \mathbb{R}^2$ and $\Omega = \overline{B\big(0, \frac{1}{3}L\big)} \subset \mathbb{R}^3$. For $\alpha \in A$, define the vertical line

$$\ell_\alpha := \{(x, y, z) : (x, y) = \alpha\}.$$

If $\alpha = (\alpha_1, \alpha_2) \in A$ and $\boldsymbol{a} = (a_1, a_2, a_3) \in \Omega$, note $(\alpha_1 - a_1, \alpha_2 - a_2) \in B$ and

$$\ell_\alpha \cap (\boldsymbol{a} + \Gamma) = (\alpha_1,\ \alpha_2,\ a_3 + \gamma(\alpha_1 - a_1, \alpha_2 - a_2))$$

is a singleton. Thus, we can define the two-parameter family of mappings $\{\Phi_\alpha(\,\cdot\,)\}_{\alpha \in A}$, $\Phi_\alpha : \Omega \to \mathbb{R}$ by

$$\Phi_\alpha(\boldsymbol{a}) = a_3 + \gamma(\alpha_1 - a_1, \alpha_2 - a_2). \tag{2-27}$$

The following lemma states that this family of maps satisfy the transversality condition of (1-13) when Γ is a surface of revolution of this form.

Lemma 2.7 (surface maps are transversal). *Let $\Gamma = \{(t, \gamma(t)) : t \in B\} = \{(t, f(|t|)) : t \in B\}$ be a surface of revolution with $f : \mathbb{R} \to \mathbb{R}$, $\gamma : \mathbb{R}^2 \to \mathbb{R}$ as defined above so that (2-26) holds on $B = \overline{B(0, L)}$. With the notation above, the associated family of projections $\{\Phi_\alpha : \Omega \to \mathbb{R} : \alpha \in A\}$ is 1-transversal in the sense of Definition 1.3.*

While the proof of Lemma 2.7 is similar to its 2-dimensional analogue, Lemma 2.6, there is a new layer of complexity that arises. In the 2-dimensional case, in which Γ was a curve and the graph of a real-valued function, the intersection set $(\boldsymbol{a} + \Gamma) \cap (\boldsymbol{b} + \Gamma)$ consisted of at most one point. Denoting this point by $\boldsymbol{x} = (x_1, x_2)$ (when it exists) and setting $H(\lambda) := |\Phi_\lambda(\boldsymbol{a}) - \Phi_\lambda(\boldsymbol{b})|$, with Φ_α as in (2-6), we saw that $H(x_1) = 0$ and observed that H grows at a linear rate in a neighborhood of x_1. In the 3-dimensional case, in which Γ is a surface, the set $(\boldsymbol{a} + \Gamma) \cap (\boldsymbol{b} + \Gamma)$ may consists of many points. Here, we show that the function $H(\lambda)$, now with Φ_α as in (2-27), obeys a similar linear growth condition along horizontal lines. We now prove Lemma 2.7 using the set-up above, and we begin with a few simplifying reductions.

Proof. By rescaling in the z-axis, we may assume that all the first partial derivatives of γ are bounded by 1. For distinct $\boldsymbol{a}, \boldsymbol{b} \in \Omega$, our aim is to verify that

$$|\{\lambda \in A : |\Phi_\lambda(\boldsymbol{a}) - \Phi_\lambda(\boldsymbol{b})| \leq \delta\}| \lesssim \frac{\delta}{|\boldsymbol{a} - \boldsymbol{b}|}.$$

Translating, it is enough to consider the situation when $\boldsymbol{a} = (0, 0, 0)$. Further, since Γ is symmetric about the origin, it suffices to consider the case when $\boldsymbol{b} = (b_1, 0, b_3)$ for $b_1, b_3 \geq 0$.

After this reduction, our goal is to show that

$$|\{\lambda \in A : |\Phi_\lambda(\vec{0}) - \Phi_\lambda(\boldsymbol{b})| \leq \delta\}| \lesssim \frac{\delta}{|\boldsymbol{b}|} \tag{2-28}$$

for a universal constant independent of $\boldsymbol{b}$ and δ. To this end, fix the coordinate λ_2 and form a slice parallel to the xz-plane; we will show that

$$|\{\lambda_1 : \lambda = (\lambda_1, \lambda_2) \in A \text{ and } |\Phi_\lambda(\vec{0}) - \Phi_\lambda(\boldsymbol{b})| \leq \delta\}| \lesssim \frac{\delta}{|\boldsymbol{b}|} \tag{2-29}$$

for a universal constant independent of λ_2. Once this is completed, we may integrate the estimate with respect to λ_2 over the interval $\left[-\frac{1}{3}L, \frac{1}{3}L\right]$ and apply Fubini's theorem to recover (2-28). Note that $|\cdot|$ in (2-28) denotes the 2-dimensional Lebesgue measure and $|\cdot|$ in (2-29) denotes the 1-dimensional Lebesgue measure.

We are now working within a 2-dimensional slice of the surface and will be able to apply the results of Section 2C. Note that the slice

$$\gamma_{\lambda_2} := \Gamma \cap \{y = \lambda_2\} = \{(t_1, \lambda_2, \gamma(t_1, \lambda_2)) : (t_1, \lambda_2) \in B\}$$

forms a curve in the plane $\{y = \lambda_2\}$. Since $\boldsymbol{b} = (b_1, 0, b_3)$, the translated surface $(\Gamma + \boldsymbol{b})$ also intersects this plane in a curve

$$(\Gamma + \boldsymbol{b}) \cap \{y = \lambda_2\} = \{(s_1 + b_1, \lambda_2, \gamma(s_1, \lambda_2) + b_3) : (s_1, \lambda_2) \in B\}.$$

The key point is that this curve is merely a translate of γ_{λ_2}:

$$(\Gamma + \boldsymbol{b}) \cap \{y = \lambda_2\} = \gamma_{\lambda_2} + \boldsymbol{b}.$$

Recalling the curvature condition (2-26), we see that the curve γ_{λ_2} satisfies the simple curvature condition of Definition 2.5. Applying Lemma 2.6 (in particular, the result of (2-12)) then establishes (2-29) as desired. □

3. Energy techniques for pushforwards

We now turn to measure estimates using the energy and potential based approach of [Mattila 1990]. The key idea here will be that the energies associated to a measure μ and its pushforwards $\widetilde{\pi}_{\alpha\sharp}\mu$ are closely related. This will allow us to prove strong asymptotic lower bounds for the Favard curve lengths of neighborhoods of sets. First, we begin by proving Proposition 1.4, illustrating how transversality plays a role in the study of pushforward measures. This proposition provides a generalization of Marstrand's result on the typical dimension of projections to a nonlinear setting, and the proof provided here is similar to that which appeared in [Solomyak 1998] in the context of general metric spaces.

Recall that we have a family $\{\widetilde{\pi_\alpha} : \alpha \in A\}$ of maps into an m-dimensional space, $\dim_{\mathcal{H}} E = t \le m$, and the family of projections is m-transversal. Our goal is to show that for ψ-almost every $\alpha \in A$,

$$\dim_{\mathcal{H}} \widetilde{\pi_\alpha} E = t.$$

The primary tool will be to use that if x and y are two fixed points, then the projection operators $\widetilde{\pi_\alpha}$ will usually be able to distinguish between x and y on scale $|x-y|$. This is quantified with the distribution function.

Proof of Proposition 1.4. Suppose that E supports a Borel probability measure μ with finite τ-energy. Recall the energy of the measure μ, $I_\tau(\mu)$, is defined in (1-1) and the pushforward, $\widetilde{\pi_\alpha}_\sharp \mu$, is defined in (1-2). Averaging over the set of parameters and computing the energies of the pushforward measures, we have

$$\begin{aligned}\int_A I_\tau(\widetilde{\pi_\alpha}_\sharp\mu)\,d\psi(\alpha) &= \int_A \iint \frac{1}{|u-v|^\tau}\,d\widetilde{\pi_\alpha}_\sharp\mu(u)\,d\widetilde{\pi_\alpha}_\sharp\mu(v)\,d\psi(\alpha)\\ &= \int_A \iint \frac{1}{|\widetilde{\pi_\alpha}(x)-\widetilde{\pi_\alpha}(y)|^\tau}\,d\mu(x)\,d\mu(y)\,d\psi(\alpha)\\ &= \iiint_A \frac{1}{|\widetilde{\pi_\alpha}(x)-\widetilde{\pi_\alpha}(y)|^\tau}\,d\psi(\alpha)\,d\mu(x)\,d\mu(y)\\ &= \iint \left[\int_A \frac{|x-y|^\tau}{|\widetilde{\pi_\alpha}(x)-\widetilde{\pi_\alpha}(y)|^\tau}\,d\psi(\alpha)\right]\frac{d\mu(x)\,d\mu(y)}{|x-y|^\tau}.\end{aligned}$$

We can study the innermost integral using the transversality condition together with the distribution function:

$$\begin{aligned}\int_A \frac{|x-y|^\tau}{|\widetilde{\pi_\alpha}(x)-\widetilde{\pi_\alpha}(y)|^\tau}\,d\psi(\alpha) &= \int_0^\infty \psi\left(\left\{\alpha : \frac{|x-y|^\tau}{|\widetilde{\pi_\alpha}(x)-\widetilde{\pi_\alpha}(y)|^\tau} \ge r\right\}\right)dr\\ &= \int_0^\infty \psi(\{\alpha : |\widetilde{\pi_\alpha}(x)-\widetilde{\pi_\alpha}(y)| \le r^{-1/\tau}|x-y|\})\,dr\\ &= \tau\int_0^\infty \psi(\{\alpha : |\widetilde{\pi_\alpha}(x)-\widetilde{\pi_\alpha}(y)| \le \delta|x-y|\})\,\frac{d\delta}{\delta^{1+\tau}}.\end{aligned}$$

For a fixed $\delta_0 > 0$, the integral on $[\delta_0, \infty)$ converges: our parameter set has finite measure, and $\int_{\delta_0}^\infty (1/\delta^{1+\tau})\,d\delta$ is finite. Therefore, we only need to consider the case of $\delta \in [0, \delta_0)$; this corresponds to the set of parameters which are not able to distinguish x and y, and will have small measure due to transversality. In particular, the m-transversality of (1-12) with $s = m$ yields

$$\psi(|\{\alpha : |\widetilde{\pi_\alpha}(x)-\widetilde{\pi_\alpha}(y) \le \delta|x-y|\}) \lesssim \delta^m$$

for all $\delta \le \delta_0$, implying

$$\int_0^{\delta_0} \psi(\{\alpha : |\widetilde{\pi_\alpha}(x)-\widetilde{\pi_\alpha}(y)| \le \delta|x-y|\})\,\frac{d\delta}{\delta^{1+\tau}} \lesssim \int_0^{\delta_0} \delta^{m-\tau}\,\frac{d\delta}{\delta}.$$

This converges provided that $\tau < m$. We have now shown that, for $\tau < m$,

$$\int_A I_\tau(\widetilde{\pi_\alpha}_\sharp\mu)\,d\alpha \lesssim \iint \frac{d\mu(x)\,d\mu(y)}{|x-y|^\tau} = I_\tau(\mu) < \infty,$$

and therefore the energy $I_\tau(\widetilde{\pi_\alpha}_\sharp\mu)$ is finite for ψ-almost every α.

To finish the proof, recall that if E has positive $\mathcal{H}^t$ measure, then for any $\tau < t$ there exists a measure μ supported on E with finite τ-energy (see Frostman's lemma in [Mattila 2015]). It follows that if $\tau < t \leq m$, then the pushforward $\widetilde{\pi}_{\alpha\sharp}\mu$ will also have finite τ-energy, implying that $\widetilde{\pi}_\alpha(E)$ has Hausdorff dimension at least τ. Passing to a countable sequence τ_n converging upwards to t gives the desired result. □

For the remainder of the section, we will employ the notation of Definition 1.2. Recall that the lower derivative of the measure $\widetilde{\pi}_{\alpha\sharp}\mu$ with respect to h at the point u is defined by

$$\underline{D}(\widetilde{\pi}_{\alpha\sharp}\mu, h, u) = \liminf_{\delta\to 0} \frac{\widetilde{\pi}_{\alpha\sharp}\mu(B(u,\delta))}{h(B(u,\delta))}.$$

The upper derivative is similarly defined, taking the limit supremum. In the case that the lower and upper derivatives coincide, they will agree with the Radon–Nikodym derivative denoted by $D(\widetilde{\pi}_{\alpha\sharp}\mu, h, u)$.

Lemma 3.1 (absolute continuity of pushforwards). *Suppose that $\{\widetilde{\pi}_\alpha : \alpha \in A\}$ is an s-transversal family of maps and that ψ is a Borel measure on A. If μ is a Borel measure with compact support contained in Ω and $I_s(\mu) < \infty$, then for ψ-almost every α we have that $\widetilde{\pi}_{\alpha\sharp}\mu \ll h$ and*

$$\int_A \int_X D(\widetilde{\pi}_{\alpha\sharp}\mu, h, u)^2 \, dh(u) \, d\psi(\alpha) \lesssim I_s(\mu).$$

Proof. Consider the integral

$$\iint \underline{D}(\widetilde{\pi}_{\alpha\sharp}\mu, h, u) \, d\widetilde{\pi}_{\alpha\sharp}\mu(u) \, d\psi(\alpha).$$

Due to the joint continuity assumption for the functions $(x,\alpha) \mapsto \widetilde{\pi}_\alpha(x)$, the integrands will be measurable with respect to the appropriate measures (each of which are Borel measures). We now follow the definition of the lower derivative along with Mattila's approach:

$$\begin{aligned}
\iint \underline{D}(\widetilde{\pi}_{\alpha\sharp}\mu, h, u) \, d\widetilde{\pi}_{\alpha\sharp}\mu(u) \, d\psi(\alpha) &= \iint \liminf_{\delta\to 0} \frac{\widetilde{\pi}_{\alpha\sharp}\mu(B(u,\delta))}{h(B(u,\delta))} \, d\widetilde{\pi}_{\alpha\sharp}\mu(u) \, d\psi(\alpha) \\
&\lesssim \liminf_{\delta\to 0} \frac{1}{\delta^m} \iint \widetilde{\pi}_{\alpha\sharp}\mu(B(u,\delta)) \, d\widetilde{\pi}_{\alpha\sharp}\mu(u) \, d\psi(\alpha) \\
&= \liminf_{\delta\to 0} \frac{1}{\delta^m} \iint \mu(\widetilde{\pi}_\alpha^{-1} B(u,\delta)) \, d\widetilde{\pi}_{\alpha\sharp}\mu(u) \, d\psi(\alpha) \\
&= \liminf_{\delta\to 0} \frac{1}{\delta^m} \iint \mu\{y : \widetilde{\pi}_\alpha y \in B(u,\delta)\} \, d\widetilde{\pi}_{\alpha\sharp}\mu(u) \, d\psi(\alpha). \qquad (3\text{-}1)
\end{aligned}$$

Pushforward measures obey the identity

$$\int g \, df_\sharp \nu = \int (g \circ f) \, d\nu$$

for nonnegative Borel functions f and g and a Borel measure ν. Applying this to the function

$$g(u) := \mu\{y : \widetilde{\pi}_\alpha y \in B(u,\delta)\},$$

we find that

$$\begin{aligned}\iint \mu\{y : \widetilde{\pi}_\alpha y \in B(u,\delta)\}\, d\widetilde{\pi}_{\alpha\sharp}\mu(u)\, d\psi(\alpha) &= \iint \mu\{y : \widetilde{\pi}_\alpha y \in B(\widetilde{\pi}_\alpha x,\delta)\}\, d\mu(x)\, d\psi(\alpha)\\ &= \iint \psi\{\alpha : \widetilde{\pi}_\alpha y \in B(\widetilde{\pi}_\alpha x,\delta)\}\, d\mu(x)\, d\mu(y)\\ &= \iint \psi(\{\alpha : \operatorname{dist}(\widetilde{\pi}_\alpha x, \widetilde{\pi}_\alpha y) \le \delta\})\, d\mu(x)\, d\mu(y). \end{aligned}\tag{3-2}$$

Combining (3-1) and (3-2), we get

$$\iint \underline{D}(\widetilde{\pi}_{\alpha\sharp}\mu, h, u)\, d\widetilde{\pi}_{\alpha\sharp}\mu(u)\, d\alpha \lesssim \liminf_{\delta\to 0} \iint \frac{\psi(\{\alpha : |\widetilde{\pi}_\alpha x - \widetilde{\pi}_\alpha y| \le \delta\})}{\delta^m}\, d\mu(x)\, d\mu(y). \tag{3-3}$$

We are now ready to apply the s-transversality condition. Since

$$\psi(\{\alpha : |\widetilde{\pi}_\alpha(x) - \widetilde{\pi}_\alpha(y)| \le \delta\}) \lesssim \frac{\delta^m}{|x-y|^s},$$

we find that

$$\iint \underline{D}(\widetilde{\pi}_{\alpha\sharp}\mu, h, u)\, d\widetilde{\pi}_{\alpha\sharp}\mu(u)\, d\psi(\alpha) \lesssim \iint \frac{d\mu(x)\, d\mu(y)}{|x-y|^s} = I_s(\mu). \tag{3-4}$$

Since $I_s(\mu) < \infty$, we conclude that for ψ-almost every α, the lower derivative $\underline{D}(\widetilde{\pi}_{\alpha\sharp}\mu, h, u)$ is finite for $\widetilde{\pi}_{\alpha\sharp}\mu$-a.e. $u \in X$. Following [Mattila 1995, Theorem 2.12], this implies that $\widetilde{\pi}_{\alpha\sharp}\mu \ll h$ for all such parameters, in which case $D(\widetilde{\pi}_{\alpha\sharp}\mu, h, u)$ exists for $\widetilde{\pi}_{\alpha\sharp}\mu$-a.e. $u \in X$. Finally, we can use Fubini's theorem to conclude that

$$\int_X D(\widetilde{\pi}_{\alpha\sharp}\mu, h, u)^2\, dh(u) = \int_X D(\widetilde{\pi}_{\alpha\sharp}\mu, h, u)\, d\widetilde{\pi}_{\alpha\sharp}\mu(u). \tag{3-5}$$

The combination of (3-5) with the estimate (3-4) establishes the desired result. □

The next result follows from Lemma 3.1 and, in essence, states that nonlinear variants of Favard length are controlled from below by the energy of any nice measure placed on the set.

Lemma 3.2 (lower bound on average projection length). *Suppose that $\{\widetilde{\pi}_\alpha : \alpha \in A\}$ is an s-transversal family of maps with a Borel probability measure ψ on A. If μ is a Borel probability measure supported on a compact set $F \subseteq \Omega$, then*

$$\int_A (h(\widetilde{\pi}_\alpha F))^{-1}\, d\psi(\alpha) \lesssim I_s(\mu)$$

and

$$\frac{1}{I_s(\mu)} \lesssim \int_A h(\widetilde{\pi}_\alpha F)\, d\psi(\alpha). \tag{3-6}$$

This is an analogue of [Mattila 1990, Theorem 3.2]. The proof relies on Lemma 3.1.

Proof. Since $F \subseteq \widetilde{\pi}_\alpha^{-1}(\widetilde{\pi}_\alpha F)$ and μ is a probability measure, we can apply the definition of the pushforward to conclude that

$$1 = \widetilde{\pi}_{\alpha\sharp}\mu(\widetilde{\pi}_\alpha F)^2 = \left(\int_{\widetilde{\pi}_\alpha F} D(\widetilde{\pi}_{\alpha\sharp}\mu, h, u)\, dh(u)\right)^2.$$

Invoking the Cauchy–Schwarz inequality,

$$1 \le h(\widetilde{\pi}_\alpha F) \int_{\widetilde{\pi}_\alpha F} D(\widetilde{\pi}_{\alpha\sharp}\mu, h, u)^2 \, dh(u)$$

for all $\alpha \in A$. After dividing both sides by $h(\widetilde{\pi}_\alpha F)$, integrating in ψ, and invoking Lemma 3.1, we have

$$\begin{aligned}\int_A (h(\widetilde{\pi}_\alpha F))^{-1} \, d\psi(\alpha) &\le \int_A \int_{\widetilde{\pi}_\alpha F} D(\widetilde{\pi}_{\alpha\sharp}\mu, h, u)^2 \, dh(u) \, d\psi(\alpha) \\ &\le \int_A \int_X D(\widetilde{\pi}_{\alpha\sharp}\mu, h, u)^2 \, dh(u) \, d\psi(\alpha) \lesssim I_s(\mu),\end{aligned}$$

thus establishing the first inequality.

For the second part of the theorem, consider the function $f(\alpha) = h(\widetilde{\pi}_\alpha F)$. Applying the Cauchy–Schwarz inequality to $1 = \int d\psi = \int f^{1/2} \cdot f^{-1/2} \, d\psi$ immediately gives the claimed result. □

In order to apply Lemma 3.2 to neighborhoods, we construct a measure with appropriate support and obtain an upper bound on the energy. The following lemma says that whenever the dimension of F is known, there is at least one auxiliary measure supported on the neighborhood $F(r)$ whose energy is easily computable. This is the final tool that we will need in order to estimate the average projection size of a neighborhood, and it comes directly from [Mattila 1990, Theorem 4.1]. We give a summary of the main idea of the construction.

Lemma 3.3 (construction of auxiliary measure). *Let $0 < s \le m$. Suppose μ is a Borel probability measure supported on a compact set $F \subset \mathbb{R}^n$ and there exists $c > 0$ so that*

$$\mu(B(x, r)) \le cr^s$$

for each $x \in \mathbb{R}^n$ and $r > 0$. Then for each $r \in (0, 1)$ there exists a probability measure ν supported in $F(2r)$ so that

$$I_s(\nu) \lesssim r^{s-m} \qquad \text{if } s < m, \tag{3-7}$$

$$I_s(\nu) \lesssim \log\left(\frac{1}{r}\right) \qquad \text{if } s = m. \tag{3-8}$$

Summary of proof. For F, μ and r as in the statement of Lemma 3.3, we can use a covering argument to find a disjoint collection of balls $\{B_i\}_{i=1}^k$, each with radius r, so that $\tau := \mu\big(\bigsqcup_{i=1}^k B_i\big) > 0$. The measure ν is then defined to be

$$\nu(A) := \frac{1}{\tau} \sum_{i=1}^k \mu(B_i) \frac{|A \cap B_i|}{|B_i|}, \tag{3-9}$$

where $|\cdot|$ denotes the n-dimensional Lebesgue measure. Note that ν is supported in $F(2r)$. The ν-measure of a ball of radius u can be bounded from above, considering the cases when $u \le r$, $r \le u \le 1$, and $1 \le u$ separately, and a computation with the distribution function (analogous to the computations in the proof of Proposition 1.4) shows that $I_m(\nu) \lesssim r^{s-m}$ when $s < m$. Further, when $s = m$, a similar computation shows that $I_m(\nu) \lesssim \log\left(\frac{1}{r}\right)$. □

With Lemmas 3.2 and 3.3 in tow, we now turn to the proof of Theorem 1.5. In this context, it will be important that the family is m-transversal, with m matching the dimension of the target set.

Proof of Theorem 1.5. Assume that $\{\widetilde{\pi_\alpha} : \alpha \in A\}$ is m-transversal. Letting F and μ be as in the hypotheses, we can use Lemma 3.3 to construct the auxiliary measure ν with computable energy. Applying the estimate (3-6) of Lemma 3.2 to ν and $F(2r)$ yields the theorem. □

4. Applications and examples

4A. ***Proving Theorems 1.6 and 1.7.*** We now turn to self-contained proofs of the applications to Favard curve length and visibility, respectively.

Proof of Theorem 1.7. In the case that Γ satisfies the simple curvature assumption of Definition 2.5, we can apply Lemma 2.6 to conclude that the curve projections associated to Fav_Γ form a 1-transversal family, and the theorem follows from Theorem 1.5. The reductions made at the beginning of Section 2C imply that establishing the theorem for this special class of Γ suffices. □

On the other hand, the visibility result requires a little bit more analysis, since transversality depends on the relative geometry of the visible set and the vantage set.

Proof of Theorem 1.6. In Lemma 2.3, we established that the family of radial projections $\{P_a : a \in A\}$ is $(n-1)$-transversal provided that the underlying probability measure ψ supported on A satisfies the tube condition with respect to E:

$$\psi(T_\delta) \lesssim \delta^{n-1}.$$

In our context, $\psi = \mathcal{H}^{n-1}$ and it suffices to show that there exists a positive $\delta > 0$ such that for any tube T_δ passing through the visible set E we have

$$\mathcal{H}^{n-1}(T_\delta \cap A) \lesssim \delta^{n-1}.$$

However, this follows immediately from the tangent plane condition: the angle between the tube T_δ and any tangent plane to A is uniformly bounded away from zero and the claim follows. Now that transversality has been established, we conclude the proof with an application of Theorem 1.5 as in the previous argument. □

A slightly more general version of Theorem 1.6 is available without separation between the vantage set A and the visible set E. The tube condition is also guaranteed upon replacing our tangent plane condition with the following slightly more technical statement: there exist $\delta_0 > 0$ and $\theta_0 \in \left(0, \frac{\pi}{2}\right)$ so that, if $a \in L_{x,y}(\delta_0) \cap A$ for distinct $x, y \in E$, then A_a meets $L_{x,y}$ at an angle of at least θ_0.

4B. ***Applications to dynamically generated sets.*** A key tool in proving Theorem 1.5 was to establish the existence of an auxiliary measure ν supported near F whose s-energy is easily computable. Lemma 3.2 then relates the average projection length to the energy. In the case of many fractal sets, we can construct the special measure ν in a geometrically motivated ad hoc manner. We now turn to the proof of Corollary 1.9.

Proof of Corollary 1.9. Set $s = m = 1$ and $r = \left(\frac{1}{4}\right)^n$. Recall $\mathcal{K}_n$ denotes the n-th generation in the construction of the four-corner Cantor set $\mathcal{K}$. We can write $\mathcal{K}_n$ as the union of 4^n squares Q_i of side length 4^{-n}, and define a probability measure on ν supported on $\mathcal{K}_n$ by

$$\nu(A) = \sum_{i=1}^{4^n} \frac{|A \cap Q_i|}{\left(\frac{1}{4}\right)^n}.$$

This is the equidistributed measure on $\mathcal{K}_n$ (and can be compared to the constructed measure of Lemma 3.3 when μ denotes the 1-Hausdorff measure restricted to $\mathcal{K}$).

Observe $\nu(\mathcal{K}_n) = 1$ and

$$\nu(B(x, u)) \sim \begin{cases} u^2/r & \text{for } u \le r, \\ u & \text{for } u \ge r, \\ 1 & \text{for } u \ge 1. \end{cases}$$

A direct estimate of the energy integral leads to

$$I_1(\nu) \sim \log\left(\frac{1}{r}\right) \sim n. \tag{4-1}$$

Next, as we have already established in Lemma 2.6 that the curve projections which lead to Fav_Γ are a 1-transversal family under our simple curvature assumption, we can apply Lemma 3.2 to conclude that

$$\frac{1}{I_1(\nu)} \lesssim \int_{\mathbb{R}} |\Phi_\alpha(\mathcal{K}_n)|\, d\alpha. \tag{4-2}$$

Combining (4-1) and (4-2) completes the proof of Corollary 1.9. □

It is worth emphasizing that the main point here is the existence of the measure ν with easily bounded energy at the appropriate dimension. As such, these techniques apply to a much broader class of fractal sets at dimension 1; whenever we can have a piece-counting argument that gives a sharp estimate for $I_1(\nu)$, we will get a similar bound. This is frequently the case for fractals that are generated by an iterated function system, including $\mathcal{K}_n$.

Next, we give the corresponding applications for visibility:

Proof of Corollary 1.8. Since no tangent line to the curve Γ passes through the compact set $[0, 1]^2$, there is a positive distance between any tangent line to Γ and $\mathcal{K}_n$. This is the 2-dimensional version of the nontangency assumption of Theorem 1.6 and thus the family of radial projections $\{P_a : a \in \Gamma\}$ is 1-transversal. Again taking ν to be the equidistributed measure on $\mathcal{K}_n$, the corollary now follows from Lemma 3.2 and the estimate (4-1). □

4C. *Projections without transversality.* In each of the cases handled above, a notion of transversality is used to show that the set of parameters which cannot distinguish two nearby points on an appropriate scale is rather small. One may ask whether such a condition is necessary. In the following examples, we explore what can happen when transversality is absent.

Example 4.1 (asymptotic Fav_Γ that decays too fast). Suppose that the curve Γ is x-axis in $\mathbb{R}^2$, suppose F is a horizontal line segment, and consider the curve projections Φ_α of Section 2C. Then Section 2C fails.

Recalling (1-8), we see that

$$\mathrm{Fav}_\Gamma(F(r)) \sim 2r.$$

This tends to zero much more rapidly than $(\log r^{-1})^{-1}$.

Our next example illustrates that Favard curve length does not necessarily detect rectifiability without a transversality assumption. In particular, without a curvature assumption, it is possible to have a purely unrectifiable set with positive and finite Hausdorff 1-measure, which has strictly positive Favard curve length.

Example 4.2 (a lower bound that does not decay). Suppose that Γ is a straight line in $\mathbb{R}^2$ passing through the origin with slope $\frac{1}{2}$ $\left(\text{or angle } \theta = \arctan \frac{1}{2}\right)$ and that F is the four-corner Cantor set. Consider the curve projections Φ_α of Section 2C. Then for all α so that Φ_α is defined on $\mathcal{K}$, the projection $\Phi_\alpha \mathcal{K}$ is an interval with length comparable to 1.

To see this, consider the first generation of the four-corner Cantor set $\mathcal{K}$ and its four constituent squares. Each square projects to an interval. Since the line has slope $\frac{1}{2}$, the points $\left(\frac{1}{4}, 0\right)$ and $\left(\frac{3}{4}, \frac{1}{4}\right)$ project to the same position within L_α. Similarly, $\left(0, \frac{1}{4}\right)$ and $\left(1, \frac{3}{4}\right)$ share a projection and so do $\left(\frac{1}{4}, \frac{3}{4}\right)$ and $\left(\frac{3}{4}, 1\right)$. Therefore, the projection of the bottom right square is a segment connecting $\widetilde{\pi_\alpha}(1, 0)$ and $\widetilde{\pi_\alpha}\left(\frac{3}{4}, \frac{1}{4}\right)$; the projection of the lower left square is a segment connecting $\widetilde{\pi_\alpha}\left(\frac{1}{4}, 0\right)$ and $\widetilde{\pi_\alpha}\left(0, \frac{1}{4}\right)$, and so on. The four intervals found in this manner only meet at their endpoints, and their union is an interval with length greater than 1. Finally, an application of self-similarity shows that this argument works for the second generation of the Cantor set as well; this extends to all subsequent generations and $\mathcal{K}$ itself.

As a final example, we see what happens for visibility when we do not assume the tube condition.

Example 4.3 (coplanar sets lack the tube condition). Suppose A and E are as in Theorem 1.6 so that A is a smooth $(n-1)$-dimensional surface, E has positive s-dimensional Hausdorff measure, and $|a-e| \lesssim 1$ for each $a \in A$ and $e \in E$. Moreover, assume that A and E are both subsets of the same hyperplane in $\mathbb{R}^n$. Consider the radial projections P_a of (1-4). Then the lower bounds of Theorem 1.6 fail when $s > n-2$.

For A and E in $\mathbb{R}^n$ and $a \in A$, the radial projection $P_a(E)$ is a set of Hausdorff dimension at most $n-1$. Embedding A and E in the same hyperplane guarantees that $P_a(E)$ is a set of Hausdorff dimension at most $n-2$. As such, it can be covered by $C\left(\frac{1}{r}\right)^{n-2}$ balls of radius r for some C. Since the $(n-1)$-dimensional measure of a ball is of order r^{n-1}, we conclude that the $(n-1)$-dimensional Hausdorff measure restricted to S^{n-1} is bounded by $|P_a(E(r))| \lesssim r$. Since $r \ll \log\left(\frac{1}{r}\right)^{-1}$ and $r \ll r^{n-1-s}$ whenever $n-2 < s$ and r is sufficiently small, both the first and second estimates of Theorem 1.6 fail in this regime.

To see what goes awry in Example 4.3, note that the tube $L_{x,y}(\delta)$ for distinct $x, y \in E(r)$ intersects A in a set of measure $\delta^{n-2} \gg \delta^{n-1}$ and the upper bound required by the tube condition in (2-2) fails. In this case, $P_a(E)$ for $a \in A$ fails to differentiate the points of E.

As an explicit example of what fails, consider the case $n = 2$. When A and E are contained in the same line, $P_a(E)$ consists of at most two points for any $a \in A$. This means that the projections P_a cannot differentiate points in E.

References

[Bateman and Volberg 2010] M. Bateman and A. Volberg, "An estimate from below for the Buffon needle probability of the four-corner Cantor set", *Math. Res. Lett.* **17**:5 (2010), 959–967. MR Zbl

[Besicovitch 1939] A. S. Besicovitch, "On the fundamental geometrical properties of linearly measurable plane sets of points, III", *Math. Ann.* **116**:1 (1939), 349–357. MR Zbl

[Bond and Volberg 2011] M. Bond and A. Volberg, "Circular Favard length of the four-corner Cantor set", *J. Geom. Anal.* **21**:1 (2011), 40–55. MR Zbl

[Bond et al. 2016] M. Bond, I. Łaba, and J. Zahl, "Quantitative visibility estimates for unrectifiable sets in the plane", *Trans. Amer. Math. Soc.* **368**:8 (2016), 5475–5513. MR Zbl

[Bongers 2019] T. Bongers, "Geometric bounds for Favard length", *Proc. Amer. Math. Soc.* **147**:4 (2019), 1447–1452. MR Zbl

[Bourgain 2010] J. Bourgain, "The discretized sum-product and projection theorems", *J. Anal. Math.* **112** (2010), 193–236. MR Zbl

[Cladek et al. 2022] L. Cladek, B. Davey, and K. Taylor, "Upper and lower bounds on the rate of decay of the Favard curve length for the four-corner Cantor set", *Indiana Univ. Math. J.* **71**:3 (2022), 1003–1025. MR Zbl

[Davey and Taylor 2022] B. Davey and K. Taylor, "A quantification of a Besicovitch non-linear projection theorem via multiscale analysis", *J. Geom. Anal.* **32**:4 (2022), art. id. 138. MR Zbl

[Hovila et al. 2012] R. Hovila, E. Järvenpää, M. Järvenpää, and F. Ledrappier, "Besicovitch–Federer projection theorem and geodesic flows on Riemann surfaces", *Geom. Dedicata* **161** (2012), 51–61. MR Zbl

[Łaba 2015] I. Łaba, "Recent progress on Favard length estimates for planar Cantor sets", pp. 117–145 in *Operator-related function theory and time-frequency analysis*, edited by K. Gröchenig et al., Abel Symp. **9**, Springer, 2015. MR Zbl

[Marstrand 1954] J. M. Marstrand, "Some fundamental geometrical properties of plane sets of fractional dimensions", *Proc. London Math. Soc.* (3) **4** (1954), 257–302. MR Zbl

[Mattila 1990] P. Mattila, "Orthogonal projections, Riesz capacities, and Minkowski content", *Indiana Univ. Math. J.* **39**:1 (1990), 185–198. MR Zbl

[Mattila 1995] P. Mattila, *Geometry of sets and measures in Euclidean spaces: fractals and rectifiability*, Cambridge Studies in Advanced Mathematics **44**, Cambridge University Press, 1995. MR Zbl

[Mattila 2015] P. Mattila, *Fourier analysis and Hausdorff dimension*, Cambridge Studies in Advanced Mathematics **150**, Cambridge University Press, 2015. MR Zbl

[Nazarov et al. 2010] F. Nazarov, Y. Peres, and A. Volberg, "The power law for the Buffon needle probability of the four-corner Cantor set", *Algebra i Analiz* **22**:1 (2010), 82–97. In Russian; translated in *St. Petersburg Math J.* **22**:1 (2011), 61–72. MR Zbl

[Peres and Schlag 2000] Y. Peres and W. Schlag, "Smoothness of projections, Bernoulli convolutions, and the dimension of exceptions", *Duke Math. J.* **102**:2 (2000), 193–251. MR Zbl

[Pollicott and Simon 1995] M. Pollicott and K. Simon, "The Hausdorff dimension of λ-expansions with deleted digits", *Trans. Amer. Math. Soc.* **347**:3 (1995), 967–983. MR Zbl

[Shmerkin 2020] P. Shmerkin, "A nonlinear version of Bourgain's projection theorem", preprint, 2020. arXiv 2003.01636

[Simon and Solomyak 2006/07] K. Simon and B. Solomyak, "Visibility for self-similar sets of dimension one in the plane", *Real Anal. Exchange* **32**:1 (2006/07), 67–78. MR Zbl

[Simon and Taylor 2020] K. Simon and K. Taylor, "Interior of sums of planar sets and curves", *Math. Proc. Cambridge Philos. Soc.* **168**:1 (2020), 119–148. MR Zbl

[Simon and Taylor 2022] K. Simon and K. Taylor, "Dimension and measure of sums of planar sets and curves", *Mathematika* **68**:4 (2022), 1364–1392. MR

[Solomyak 1995] B. Solomyak, "On the random series $\sum \pm\lambda^n$ (an Erdős problem)", *Ann. of Math.* (2) **142**:3 (1995), 611–625. MR Zbl

[Solomyak 1998] B. Solomyak, "Measure and dimension for some fractal families", *Math. Proc. Cambridge Philos. Soc.* **124**:3 (1998), 531–546. MR Zbl

[Tao 2009] T. Tao, "A quantitative version of the Besicovitch projection theorem via multiscale analysis", *Proc. Lond. Math. Soc.* (3) **98**:3 (2009), 559–584. MR Zbl

[Tolsa 2002] X. Tolsa, "On the analytic capacity γ_+", *Indiana Univ. Math. J.* **51**:2 (2002), 317–343. MR Zbl

Received 17 May 2021. Revised 4 Feb 2022. Accepted 22 Mar 2022.

ROSEMARIE BONGERS: bongers@math.harvard.edu
Department of Mathematics, Harvard University, Cambridge, MA, United States

KRYSTAL TAYLOR: taylor.2952@osu.edu
Department of Mathematics, The Ohio State University, Columbus, OH, United States

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at msp.org/apde.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in APDE are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

ANALYSIS & PDE

Volume 16 No. 1 2023