DENSITY OF RATIONAL POINTS ON MANIFOLDS AND DIOPHANTINE APPROXIMATION ON HYPERSURFACES

SHUNTARO YAMAGISHI

ABSTRACT. In this article, we establish an analogue of the dimension growth conjecture, which is regarding the density of rational points on projective varieties, for compact submanifolds of $\mathbb{R}^n$ with non-vanishing curvature. We also establish the convergence theory for the set of simultaneously $\psi$-approximable points lying on a generic hypersurface, thereby settling the generalized Baker-Schmidt problem in the simultaneous setting for generic hypersurfaces. These results are obtained as consequences of an optimal upper bound for the density of rational points near manifolds of the form $\{(x, f(x)) \in \mathbb{R}^{d+1} : x \in B_\nu(x_0)\}$ with non-zero Hessian matrix of $f$ at $x_0$ and $\nu > 0$ sufficiently small.

CONTENTS

1. Introduction 1
2. Notations 7
3. Preliminaries 8
4. Proof of Theorem 1.8 15
5. Proof of Proposition 4.4 22
Appendix A. Differential Geometry by Matthew Beckett 36
References 39

1. INTRODUCTION

The study of rational points on algebraic varieties and Diophantine approximation on manifolds hold prominent places in number theory. In this article, we establish results that are relevant to both of these important areas.

1.1. The dimension growth conjecture. The following conjecture, raised as a question by Serre [12, p. 27], has been greatly influential in number theory and algebraic geometry, particularly in the study of the arithmetic of higher dimensional projective varieties.

Date: February 25, 2022.
2010 Mathematics Subject Classification. 11J83 (11J13, 11J25).
Conjecture. Let $X \subseteq \mathbb{P}^{n-1}_\mathbb{Q}$ be an integral projective variety of degree $d \geq 2$ defined over $\mathbb{Q}$. Let

$$N_X(B) = \# \{ a \in \mathbb{Z}_{\text{prim}}^n : [a_1 : \cdots : a_n] \in X, |a| \leq B \}.$$

Then

$$N_X(B) \ll B^{\dim X + \varepsilon}$$

for any $\varepsilon > 0$.

This conjecture, known as the dimension growth conjecture, is now a theorem due to Salberger [11], while a large portion of it was already established in [4] and [7], and the factor $B^{\varepsilon}$ was removed in [5] when the degree $d \geq 5$ (we refer the reader to [5] for a more complete introduction to this topic).

We establish an analogue of the dimension growth conjecture for a natural class of submanfolds of $\mathbb{R}^n$.

Theorem 1.1. Let $M \subseteq \mathbb{R}^n$ be a compact smooth submanifold of $\mathbb{R}^n$ with non-vanishing curvature. Let

$$N_M(Q) = \# \{ (a, q) \in \mathbb{Z}^n \times \mathbb{N} : a/q \in M, q \leq Q \}.$$

Then

$$N_M(Q) \ll Q^{\dim M + \varepsilon}$$

for any $\varepsilon > 0$.

Remark 1.2. The precise notion of curvature in Theorem 1.1 is the Riemann curvature tensor, which is generally considered to be the most natural notion of curvature for Riemannian manifolds (submanifolds of $\mathbb{R}^n$ are canonically Riemannian) in differential geometry; by non-vanishing curvature, we mean that the Riemann curvature tensor does not vanish at $p$ for all $p \in M$. In fact, we prove Theorem 1.1 for a more general class of submanifolds of $\mathbb{R}^n$, namely compact smooth submanifolds of $\mathbb{R}^n$ with nowhere vanishing second fundamental form. The second fundamental form can be thought of as a measure of how “curved” the embedding is. As we shall see in Appendix A, the non-vanishing of any of the Riemann curvature tensor, Ricci, scalar or mean curvature at $p$ implies the non-vanishing of the second fundamental form at $p$.

Since $M$ in consideration is compact, by the implicit function theorem we may deduce Theorem 1.1 from the following local version of the result.

Theorem 1.3. Let $n > d \geq 1$, $f_1, \ldots, f_{n-d} \in C^\infty(\mathbb{R}^d)$, $x_0 \in \mathbb{R}^d$, $\nu > 0$ and

$$M = \{ (x, f_1(x), \ldots, f_{n-d}(x)) \in \mathbb{R}^n : x \in B_\nu(x_0) \}.$$

1By submanifolds, we mean embedded submanifolds.
Suppose there exists \(1 \leq i \leq n - d\) such that \(H_{f_i}(x_0)\), the Hessian matrix of \(f_i\) at \(x_0\), is a non-zero matrix and \(\nu > 0\) is sufficiently small. Then
\[
N_M(Q) \ll Q^{d+\varepsilon}
\]
for any \(\varepsilon > 0\).

**Remark 1.4.** As we shall see in Appendix A for \(M\) as in the statement of Theorem 1.3 the non-vanishing of the second fundamental form at \(p = (x, f_1(x), \ldots, f_{n-d}(x))\) is equivalent to the existence of \(1 \leq i \leq n - d\) such that \(H_{f_i}(x)\) is a non-zero matrix.

**Remark 1.5.** Given an immersion with compact domain, there exists a finite open cover of the domain such that the restriction of the immersion to each element of the cover is an embedding.

In order to establish the conclusion of Theorem 1.3 work prior to this article [9, Theorem 4] required the existence of an invertible linear combination of \(H_{f_1}(x_0), \ldots, H_{f_{d-r}}(x_0)\), while we only require one of \(H_{f_i}(x_0)\) to be a non-zero matrix.

The estimates in Theorems 1.1 and 1.3 are best possible (up to a factor of \(Q^\varepsilon\)) as the following example shows. Let \(\nu > 0\) and \(M = \{(x_1, x_2^2) \in \mathbb{R}^2 : x_1 \in B_\nu(0)\}\). Then
\[
N_M(Q) \gg \sum_{1 \leq q \leq \sqrt{Q}} \sum_{1 \leq a \leq q} 1 \gg Q^{\dim M}.
\]

### 1.2. Diophantine approximation on hypersurfaces.

Given a hypersurface \(V \subseteq \mathbb{R}^n\), i.e. \(\dim V = n - 1\), let us define \(V^*\) to be the set of points \(p \in V\) where the second fundamental form at \(p\) vanishes. We say \(V\) is **generic** if \(V^*\) is countable. Let \(f \in C^\infty(\mathbb{R}^{n-1})\), \(x_0 \in \mathbb{R}^{n-1}\), \(\nu > 0\) and \(V = \{(x, f(x)) \in \mathbb{R}^n : x \in B_\nu(x_0)\}\). Then
\[
V^* = \left\{(x, f(x)) \in \mathbb{R}^n : x \in B_\nu(x_0), \frac{\partial^2 f}{\partial x_i \partial x_j}(x) = 0 \quad (1 \leq i \leq j \leq n - 1)\right\}.
\]

It follows that
\[
\pi(V^*) = \bigcap_{1 \leq i < j \leq n-1} \left\{x \in B_\nu(x_0) : \frac{\partial^2 f}{\partial x_i \partial x_j}(x) = 0 \right\},
\]
where \(\pi : \mathbb{R}^n \to \mathbb{R}^{n-1}\) is the projection onto the first \(n - 1\) coordinates, is an intersection of \(n(n-1)/2\) zero sets of smooth functions. Therefore, we expect \(V^*\) to be countable for a typical \(f\), and the assumption that a hypersurface is generic is a natural assumption.

Let \(H^s\) denote the \(s\)-dimensional Hausdorff measure. Given a decreasing function \(\psi : \mathbb{N} \to \mathbb{R}_{>0}\), a point \(z = (z_1, \ldots, z_n) \in \mathbb{R}^n\) is called **simultaneously**
ψ-approximable if there are infinitely many \( q \in \mathbb{N} \) such that
\[
\max_{1 \leq i \leq n} \|qz_i\| < \psi(q).
\]
We denote by \( \mathcal{S}_n(\psi) \) the set of simultaneously ψ-approximable points in \( \mathbb{R}^n \). We establish the convergence theory for the set of simultaneously ψ-approximable points lying on a generic hypersurface.

**Theorem 1.6.** Let \( n > 1 \), \((n - 1)/2 < s < n - 1\) and \( V \subseteq \mathbb{R}^n \) a generic smooth hypersurface. Let \( \psi : \mathbb{N} \to \mathbb{R}_{>0} \) be a decreasing function. Then
\[
\mathcal{H}^s(\mathcal{S}_n(\psi) \cap V) = 0 \quad \text{if} \quad \sum_{q=1}^{\infty} q^n \left( \frac{\psi(q)}{q} \right)^{s+1} < \infty.
\]

In order to establish the conclusion of Theorem 1.6, work prior to this article [9, Theorem 5] required the non-vanishing of the Gaussian curvature everywhere except on a set of zero \( s \)-dimensional Hausdorff measure. For \( V = \{(x, f(x)) \in \mathbb{R}^n : x \in B_r(x_0)\} \), this condition means
\[
\mathcal{H}^s(\{(x, f(x)) \in \mathbb{R}^n : x \in B_r(x_0), \det H_f(x) = 0\}) = 0;
\]
however, in general we expect this value to be infinity for \( s < n - 2 \) (if the regular value theorem applies, then the set in consideration is an \((n - 2)\)-dimensional manifold and its \( s \)-dimensional Hausdorff measure is infinity for \( s < n - 2 \)). On the other hand, we only require
\[
\mathcal{H}^s(V \cap V^*) = 0
\]
or \( V^* \) to be countable (the \( s \)-dimensional Hausdorff measure of a countable set is zero), which as explained above is a natural assumption.

In contrast to the convergence case, the divergence case has been established in greater generality, for non-degenerate analytic manifolds, in [1]. Since a generic analytic hypersurface \( V \) is non-degenerate, Theorem 1.6 complements the divergence case\(^2\) and settles the generalized Baker-Schmidt problem in the simultaneous setting for generic hypersurfaces (we refer the reader to [1] and [2] for a more complete introduction to this topic).

**Corollary 1.7.** Let \( n > 1 \), \((n - 1)/2 < s < n - 1\) and \( V \subseteq \mathbb{R}^n \) a generic analytic hypersurface. Let \( \psi : \mathbb{N} \to \mathbb{R}_{>0} \) be a decreasing function. Then
\[
\mathcal{H}^s(\mathcal{S}_n(\psi) \cap V) = \begin{cases} 0 & \text{if} \quad \sum_{q=1}^{\infty} q^n \left( \frac{\psi(q)}{q} \right)^{s+1} < \infty, \\ \infty & \text{if} \quad \sum_{q=1}^{\infty} q^n \left( \frac{\psi(q)}{q} \right)^{s+1} = \infty. \end{cases}
\]

\(^2\)In the statement of [1, Theorem 2.5], there is an additional technical condition \( q \psi(q)^m \to \infty \) as \( q \to \infty \). This condition can be removed if the corresponding convergence case is proved as in Corollary 1.7.
1.3. Density of rational points near hypersurfaces. Let $d \geq 1$, $\mathcal{D} \subseteq \mathbb{R}^d$ an open set, $x_0 \in \mathcal{D}$ and $\nu > 0$. Given $f \in C^\infty(\mathcal{D})$, $0 \leq \delta \leq 1/2$ and $Q \geq 1$, let us define

$$N_f(Q, \delta) = \sum_{1 \leq q \leq Q} \sum_{a \in \mathbb{Z}^d \atop \|qf(a/q)\| \leq \delta} 1_{B_\nu(x_0)}\left(\frac{a}{q}\right),$$

where $\|\cdot\|$ denotes the distance to the nearest integer. Since $\|x\| \leq 1/2$ ($x \in \mathbb{R}$), we only consider $0 \leq \delta \leq 1/2$. It is clear that $N_f(Q, \delta)$ is the number of rational points, with bounded denominators, near the manifold $\{(x, f(x)) \in \mathbb{R}^{d+1} : x \in B_\nu(x_0)\}$. A trivial upper bound for $N_f(Q, \delta)$ is given by

$$N_f(Q, \delta) \ll Q^{d+1},$$

while a probabilistic heuristic suggests

$$\delta Q^{d+1} \ll N_f(Q, \delta) \ll \delta Q^{d+1},$$

which we know not to hold in complete generality. There has been a significant progress\footnote{The results mentioned in this section do not require $f$ to be smooth. We refer the reader to the respective articles for the precise conditions.} in estimating this quantity $N_f(Q, \delta)$. The starting point was the following estimate, which holds when $d = 1$, $\frac{\partial^2 f}{\partial x_1^2}(x_0) \neq 0$ and $\nu$ is sufficiently small, due to Huxley [10]

$$N_f(Q, \delta) \ll \delta^{1-\varepsilon}Q^2 + Q(\log Q) \quad (1.1)$$

for any $\varepsilon > 0$. This estimate was improved by Vaughan and Velani [13] to

$$N_f(Q, \delta) \ll \delta Q^2 + Q^{1+\varepsilon}. \quad (1.2)$$

A higher dimensional generalization, which holds when $d > 1$, $H_f(x_0)$ is invertible and $\nu$ is sufficiently small, was first obtained by Beresnevich, Vaughan, Velani and Zorin [3]

$$N_f(Q, \delta) \ll \delta Q^{d+1} + Q^{d+\frac{2d}{d+2}}(\log Q)^\frac{2d}{d+2}. \quad (1.3)$$

In a major achievement by Huang [9], this estimate was improved to

$$N_f(Q, \delta) \ll \begin{cases} \delta Q^3 + Q^2 \exp(c\sqrt{\log Q}) & \text{if } d = 2, \\ \delta Q^{d+1} + Q^d(\log Q)^c & \text{if } d > 2, \end{cases} \quad (1.3)$$

where $c > 0$ is a constant depending only on $d$, $x_0$, $\nu$ and $f$. In this article, we relax the condition that $H_f(x_0)$ is invertible to only requiring it to be a non-zero matrix.
Theorem 1.8. Let $f \in C^\infty(\mathcal{O})$. Suppose $H_f(x_0)$ is a non-zero matrix and $\nu > 0$ is sufficiently small. Then

$$N_f(Q, \delta) \ll \delta Q^{d+1} + Q^{d+\varepsilon}$$

for any $\varepsilon > 0$.

Our main results Theorems 1.3 and 1.6 follow from Theorem 1.8. If we let $M$ be as in the statement of Theorem 1.3, then

$$N_M(Q) \ll \min_{1 \leq i \leq d} N_{f_i}(Q, \delta),$$

and the conclusion of Theorem 1.3 follows from Theorem 1.8 with $\delta = 0$. We omit the deduction of Theorem 1.6 from Theorem 1.8 as the argument is identical to that of [9, Theorem 5] on replacing the condition $\det H_f(x) \neq 0$ with $H_f(x)$ being non-zero in the proof.

The assumption in Theorem 1.8 is optimal in the sense that there are examples of $f \in C^\infty(\mathcal{O})$ where $H_f(x_0)$ is the zero matrix and

$$\frac{\partial^{i_1 + \cdots + i_d} f}{\partial^{i_1} x_1 \cdots \partial^{i_d} x_d}(x_0) \neq 0$$

for some $i_1, \ldots, i_d \in \mathbb{Z}_{\geq 0}$ satisfying $i_1 + \cdots + i_d > 2$ such that the inequality does not hold. As explained in [9, Example 4], one such example is the Fermat curve. Let $k > 2$ and

$$f(x_1) = (1 - x_1^k)^{\frac{1}{k}}.$$

It is clear that $\frac{\partial^2 f}{\partial x_1^2}(0) = 0$ and $\frac{\partial^k f}{\partial x_1^k}(0) \neq 0$. Then from a simple application of Taylor’s theorem we may deduce that

$$N_f(Q, \delta) \gg \sum_{1 \leq q \leq Q} \sum_{\delta q^{k-1} \leq a \leq \delta q} 1 \gg \delta^{\frac{1}{k}} Q^{2-\frac{d}{k}}$$

when $\delta \gg \frac{1}{Q}$.

Therefore, we see that the inequality in Theorem 1.8 cannot hold in this case.

We prove Theorem 1.8 by developing a method to reduce the problem to that for curves in $\mathbb{R}^2$ or surfaces in $\mathbb{R}^3$ based on whether there exists a non-zero diagonal entry of $H_f(x_0)$ or not. We then use (1.1) and (1.2) as keys inputs in the former case and (1.3) with $d = 2$ in the latter case. This is in stark contrast to the approach in [9], which treats the problem as a $d$-dimensional problem and makes use of all the $d$ variables, while only at most two variables have the same role in our approach; this is perhaps the most surprising part of our method, since the use of the $d$-dimensional stationary phase estimate is crucial in [9], and the challenge lies in dealing with the non-essential variables efficiently, that is without introducing additional factors to the final estimate. In fact, our approach gives a new proof of (1.3) when $d > 2$ (with $Q^\varepsilon$ instead of $(\log Q)^\varepsilon$). We cover notations in Section 2 and preliminaries to prove Theorem 1.8 in Section 3. An important first step in our method, in essence, is to find a
suitable function $g$, which in this context means that the Legendre transform of $g$ is linear with respect to the non-essential variables, and $\alpha > 0$ such that

$$Nf(Q, \delta) \ll Ng(Q, \alpha).$$

This technical procedure is described in Section 4, and an appropriate upper bound for the latter term is established in Section 5, which completes the proof of Theorem 1.8.

Acknowledgements. The author was supported by the NWO Veni Grant 016.Veni.192.047 while working on this article. He would like to thank Matthew Beckett for writing the appendix, Damaris Schindler for introducing him to this problem, Fabian Ziltener for all the discussions on various notions in differential geometry, and Victor Beresnevich, Raf Cluckers and Jing-Jing Huang for their helpful correspondences.

2. Notations

In this article, $|\cdot|$ denotes the $L^\infty$-norm, i.e. $|t| = \max_{1 \leq i \leq m} |t_i|$ for $t = (t_1, \ldots, t_m) \in \mathbb{R}^m$. For $\kappa > 0$, we let

$$B_\kappa(t) = (t_1 - \kappa, t_1 + \kappa) \times \cdots \times (t_m - \kappa, t_m + \kappa).$$

Given $\mathcal{U} \subseteq \mathbb{R}^m$, we denote by $\partial \mathcal{U}$ the boundary of $\mathcal{U}$, i.e. $\partial \mathcal{U} = \overline{\mathcal{U}} \setminus \mathcal{U}^o$ where $\overline{\mathcal{U}}$ is the closure of $\mathcal{U}$ and $\mathcal{U}^o$ is the interior of $\mathcal{U}$. We let $\mathbbm{1}_\mathcal{U}$ be the characteristic function of $\mathcal{U}$, i.e. $\mathbbm{1}_\mathcal{U}(t) = 1$ if $t \in \mathcal{U}$ and 0 otherwise. Given $k \in \mathbb{N}$ and $\mathcal{V} \subseteq \mathbb{R}^m$ an open set, we denote by $C^0(\mathcal{V})$, $C^k(\mathcal{V})$ and $C^\infty(\mathcal{V})$ the sets of continuous, $k$-times continuously differentiable and smooth functions defined on $\mathcal{V}$ respectively, and let

$$C^\infty_0(\mathbb{R}^m) = \{ \omega \in C^\infty(\mathbb{R}^m) : \text{supp} \omega \text{ is compact} \},$$

where

$$\text{supp} \omega = \{ t \in \mathbb{R}^m : \omega(t) \neq 0 \}.$$ 

Given $k \in \mathbb{Z}_{\geq 0}$ and $F \in C^k(\mathcal{V})$, we let

$$\|F\|_{C^k(\mathcal{V})} = \sum_{i=0}^k \sum_{i_1, \ldots, i_m \geq 0} \sup_{t \in \mathcal{V}} \left| \frac{\partial^i F}{\partial t_{i_1} \cdots \partial t_{i_m}}(t) \right|,$$

$$\nabla F(t) = \left( \frac{\partial F}{\partial t_1}(t), \ldots, \frac{\partial F}{\partial t_m}(t) \right) \quad (t \in \mathcal{V})$$

when $k \geq 1$, and

$$H_F(t) = \left[ \frac{\partial^2 F}{\partial t_i \partial t_j}(t) \right]_{1 \leq i, j \leq m} \quad (t \in \mathcal{V}).$$
when \( k \geq 2 \), which is the Hessian matrix of \( F \) at \( t \). Suppose \( r \in \mathbb{N}, \mathcal{U} \subseteq \mathbb{R}^r \) an open set and \( G \in C^k(\mathcal{U} \times \mathcal{V}) \). We denote

\[
G_t(y) = G(y, t) \quad (y \in \mathcal{U})
\]

for each \( t \in \mathcal{V} \). For any \( z \in \mathbb{R} \), we let \( e(z) = e^{2\pi iz} \). By the notation \( \Im(x) \ll \Re(x) \) or \( \Im(x) = O(\Re(x)) \), we mean that there exits a constant \( c > 0 \) such that \( |\Im(x)| < c \Re(x) \) for all \( x \) in consideration. Finally, \( \varepsilon > 0 \) will always be assumed to be sufficiently small (with respect to relevant parameters) even when it is not explicitly stated so.

3. Preliminaries

In this section, we collect various results we need to prove Theorem 1.8.

3.1. Legendre transform. Let \( r \geq 1, \ell \geq 2, \mathcal{U} \subseteq \mathbb{R}^r \) an open set and \( F \in C^\ell(\mathcal{U}) \) such that \( \nabla F \) is invertible on \( \mathcal{U} \). We define the Legendre transform \( F^* : \nabla F(\mathcal{U}) \to \mathbb{R} \) of \( F \) by

\[
F^*(u) = u \cdot (\nabla F)^{-1}(u) - (F \circ (\nabla F)^{-1})(u).
\]

It can be verified that \( F^* \in C^\ell(\nabla F(\mathcal{U})) \), \( (F^*)^* = F \) and \( \nabla F^* = (\nabla F)^{-1} \). In particular, for \( u = \nabla F(y) \) we have

\[
F^*(u) = u \cdot y - F(y) \tag{3.1}
\]

and

\[
H_{F^*}(u) = H_F(y)^{-1}. \tag{3.2}
\]

Lemma 3.1. Let \( r, m \geq 1, \ell \geq 2, \mathcal{U} \subseteq \mathbb{R}^r \) and \( \mathcal{V} \subseteq \mathbb{R}^m \) open sets, \( y_0 \in \mathcal{U} \) and \( t_0 \in \mathcal{V} \). Let \( G \in C^\ell(\mathcal{U} \times \mathcal{V}) \). Suppose \( \det G_{t_0}(y_0) \neq 0 \). Let \( \tau, \mu > 0 \) be sufficiently small. Then the following statements hold:

(i) \( \nabla G_t \) is a \( C^{\ell-1} \)-diffeomorphism on \( B_\tau(y_0) \) for each \( t \in B_\tau(t_0) \).

(ii) \( (\nabla G_t)^{-1} \) is a \( C^{\ell-1} \)-diffeomorphism on \( B_\mu(\nabla G_{t_0}(y_0)) \) for each \( t \in B_\mu(t_0) \).

(iii) Let us denote

\[
\psi(u, t) = (\psi_1(u, t), \ldots, \psi_r(u, t)) = (\nabla G_t)^{-1}(u).
\]

Then \( \psi_k \in C^{\ell-1}(B_\mu(\nabla G_{t_0}(y_0)) \times B_\mu(t_0)) \) for each \( 1 \leq k \leq r \).

Proof. The result follows as a consequence of the inverse function theorem with \( \varsigma : \mathcal{U} \times \mathcal{V} \to \mathbb{R}^{r+m} \) defined by

\[
\varsigma(y, t) = (\nabla G_t(y), t).
\]

\[\square\]

Lemma 3.2. Let \( d > r \geq 1, \ell \geq 2, \mathcal{U} \subseteq \mathbb{R}^r \) and \( \mathcal{V} \subseteq \mathbb{R}^{d-r} \) open sets, \( y_0 \in \mathcal{U} \) and \( t_0 \in \mathcal{V} \). Let \( G \in C^\ell(\mathcal{U} \times \mathcal{V}) \). Suppose \( \det G_{t_0}(y_0) \neq 0 \). Let \( \mu > 0 \) be sufficiently small. Then the following statements hold:
(i) Let us denote \( \varphi(u, t) = G_t^*(u) \). Then \( \varphi \in C^{\ell-1}(B_\mu(\nabla G_{t_0}(y_0)) \times B_\mu(t_0)) \).

(ii) \( \nabla G_t^* \) is a \( C^{\ell-1} \)-diffeomorphism on \( B_\mu(\nabla G_{t_0}(y_0)) \) for each \( t \in B_\mu(t_0) \).

(iii) We have that
\[
(G_t^*)^*(y) = G_t(y)
\]
holds for all \( y \in B_\mu(y_0) \) and \( t \in B_\mu(t_0) \).

Proof. Let us denote
\[
\psi(u, t) = (\nabla G_t)^{-1}(u).
\]
Since
\[
\varphi(u, t) = G_t^*(u) = u \cdot \psi(u, t) - G(\psi(u, t), t)
\]
and
\[
\nabla G_t^* = (\nabla G_t)^{-1},
\]
statements (i) and (ii) follow by Lemma 3.1. Let us denote
\[
\tilde{G}(y, t) = \varphi_t^*(y).
\]
By applying statement (i) now to \( \varphi \), it follows that \( \tilde{G} \in C^{\ell-2}(B_{\mu'}(y_0) \times B_{\mu'}(t_0)) \) for \( \mu' > 0 \) sufficiently small. Since
\[
\varphi_t^*(y) = (G_t^*)^*(y) = G_t(y),
\]
statement (iii) follows. \( \square \)

**Lemma 3.3.** Let \( d > r \geq 1, \ell \geq 3, \mu > 0, u_0 \in \mathbb{R}^r \) and \( z_0 \in \mathbb{R}^{d-r} \). Suppose \( \mathcal{F} \in C^{\ell+1}(B_\mu(u_0) \times B_\mu(z_0)) \) and \( \det H_{\mathcal{F}_z_0}(u_0) \neq 0 \). We define
\[
\mathcal{L}(u, z, v) = \mathcal{F}(u, v) + \sum_{1 \leq i \leq d-r} \frac{\partial \mathcal{F}}{\partial z_i}(u, v)(z_i - v_i)
\]
for \( (u, z, v) \in B_\mu(u_0) \times B_\mu(z_0) \times B_\mu(z_0) \). Let us denote
\[
g(y, z, v) = \mathcal{L}_{(z, v)}^*(y).
\]
Let \( \tau, \nu > 0 \) be sufficiently small and \( y_0 = \nabla \mathcal{F}_{z_0}(u_0) \). Then the following statements hold:

(i) \( g \in C^{\ell-1}(B_\tau(y_0) \times B_{2\nu}(z_0) \times B_{3\nu}(z_0)). \)

(ii) \( g_{(z, v)} \in C^{\ell}(B_\tau(y_0)) \) and \( \nabla g_{(z, v)} \) is a \( C^{\ell-1} \)-diffeomorphism on \( B_\tau(y_0) \) for each \( z \in B_{2\nu}(z_0) \) and \( v \in B_{3\nu}(z_0) \).

(iii) There exist \( c_-, c_+ > 0 \) such that
\[
c_- \leq |\det H_{g_{(z, v)}(y_0)}(y)| \leq c_+ \quad (y \in B_\tau(y_0), z \in B_{2\nu}(z_0), v \in B_{3\nu}(z_0)).
\]

(iv) For \( \nu \) sufficiently small with respect to \( \tau \), there exists \( \rho > 0 \) such that
\[
\text{dist}(\partial(\nabla g_{(z, v)}(B_\tau(y_0))), \partial(\nabla g_{(z, v)}(B_{2\nu}(y_0)))) \geq \rho \quad (z \in B_{2\nu}(z_0), v \in B_{3\nu}(z_0)).
\]
(v) Given \( y \in B_r(y_0) \) and \( \nu \in B_{3\nu}(z_0) \), the first degree Taylor polynomial (with respect to the \( z \) variables) of
\[
\mathcal{F}_z^*(y) - \mathcal{L}^*_{(z, \nu)}(y)
\]
centered at \( \nu \) is the zero function.

(vi) There exists \( C > 0 \) such that
\[
|\mathcal{F}_z^*(y) - \mathcal{L}^*_{(z, \nu)}(y)| \leq C|z - \nu|^2
\]
holds for all \( y \in B_r(y_0) \), \( z \in B_{2\nu}(z_0) \) and \( \nu \in B_{3\nu}(\nu_0) \).

**Proof.** Since
\[
\mathcal{L}(u, z_0, z_0) = \mathcal{F}(u, z_0),
\]
we have
\[
\det H_{\mathcal{L}_{(u, z_0)}}(u_0) = \det H_{\mathcal{F}_{z_0}}(u_0) \neq 0.
\]
Therefore, statements (i) and (ii) follow from Lemma 3.2. We also have
\[
\det H_{\mathcal{L}_{(u, z_0)}}(y_0) = \det H_{\mathcal{L}^*_{(z, \nu)}}(y_0) = \det H_{\mathcal{F}_z^*}(y_0) = \frac{1}{\det H_{\mathcal{F}_{z_0}}(u_0)} \neq 0,
\]
and statement (iii) follows by continuity of the entries of \( H_{\mathcal{F}_{(z, \nu)}}(y) \). It is easy to see that statement (iv) follows from statement (ii).

Let us denote
\[
\psi(y, z, \nu) = (\psi_1(y, z, \nu), \ldots, \psi_r(y, z, \nu)) = (\nabla \mathcal{L}_{(z, \nu)})^{-1}(y).
\]
It follows from Lemma 3.3 that there exist \( 0 < \kappa_0 < \kappa_1 < \kappa_2 < \mu/2 \) such that for all \( z, \nu \in B_{\kappa_0}(z_0) \) we have \( \nabla \mathcal{L}_{(z, \nu)} \) and \( (\nabla \mathcal{L}_{(z, \nu)})^{-1} \) are \( C^{r-1} \)-diffeomorphisms on \( B_{\kappa_1}(u_0) \) and \( B_{\kappa_2}(y_0) \) respectively, and \( \nabla \mathcal{F}_z \) and \( (\nabla \mathcal{F}_z)^{-1} \) are \( C^r \)-diffeomorphisms on \( B_{\kappa_1}(u_0) \) and \( B_{\kappa_2}(y_0) \) respectively, and
\[
\psi_k \in C^{r-1}(B_{\kappa_2}(y_0) \times B_{\kappa_0}(z_0) \times B_{\kappa_0}(z_0))
\]
for each \( 1 \leq k \leq r \). We may assume without loss of generality that \( \kappa_1 \) is sufficiently small with respect to \( \kappa_2 \) such that
\[
\nabla \mathcal{F}_{z_0}(B_{2\kappa_1}(u_0)) \subseteq B_{\kappa_2}(y_0).
\]
Let \( \tau \) and \( \nu \) be sufficiently small such that
\[
\bigcup_{z \in B_{3\nu}(z_0)} \nabla \mathcal{F}_z(B_{\kappa_1}(u_0)) \cup \nabla \mathcal{L}_{(z, \nu)}(B_{\kappa_1}(u_0)) \subseteq B_{\kappa_2}(y_0)
\]
and
\[
\bigcup_{z \in B_{3\nu}(z_0)} (\nabla \mathcal{F}_z)^{-1}(B_{\tau}(y_0)) \cup (\nabla \mathcal{L}_{(z, \nu)})^{-1}(B_{\tau}(y_0)) \subseteq B_{\kappa_1}(u_0).
\] (3.3)
Let $v \in B_{3\nu}(z_0)$. Since
$$\mathcal{F}(u, v) = \mathcal{L}(u, v, v),$$
we have
$$\nabla \mathcal{F}(u) = \nabla \mathcal{L}_{(v,v)}(u). \quad (3.4)$$
It follows from (3.3) that
$$\mathcal{F}^*_v(y) = y \cdot (\nabla \mathcal{F})^{-1}(y) - \mathcal{F}((\nabla \mathcal{F})^{-1}(y))$$
$$= y \cdot (\nabla \mathcal{L}_{(v,v)})^{-1}(y) - \mathcal{L}_{(v,v)}((\nabla \mathcal{L}_{(v,v)})^{-1}(y))$$
$$= \mathcal{L}^*_{(v,v)}(y)$$
holds for all $y \in B_{r}(y_0)$ and $v \in B_{3\nu}(z_0)$.
Let us denote
$$\xi(u, z, v) = (\xi_1(u, z, v), \ldots, \xi_r(u, z, v)) = \nabla \mathcal{L}(z, v)(u)$$
and
$$\zeta(y, z) = (\zeta_1(y, z), \ldots, \zeta_r(y, z)) = (\nabla \mathcal{F}^{-1}(y)).$$
Let $1 \leq i \leq d - r$. By taking the derivative with respect to $z_i$ of the identity
$$\psi(\xi(u, z, v), z, v) = (\nabla \mathcal{L}_{(z,v)})^{-1}(\nabla \mathcal{L}_{(z,v)}(u)) = u,$$
we obtain
$$0 = \sum_{1 \leq j \leq r} \frac{\partial \psi_k}{\partial y_j}(\xi(u, z, v), z, v) \cdot \frac{\partial \xi_j}{\partial z_i}(u, z, v) + \frac{\partial \psi_k}{\partial z_i}(\xi(u, z, v), z, v) \quad (3.5)$$
for each $1 \leq k \leq r$. It is clear that
$$\frac{\partial \mathcal{L}}{\partial z_i}(u, z, v) = \frac{\partial \mathcal{F}}{\partial z_i}(u, v), \quad (3.6)$$
$$\frac{\partial \xi_j}{\partial z_i}(u, z, v) = \frac{\partial^2 \mathcal{L}}{\partial z_i \partial u_j}(u, z, v) = \frac{\partial^2 \mathcal{F}}{\partial z_i \partial u_j}(u, v)$$
and
$$\psi(y, v, v) = (\nabla \mathcal{L}_{(v,v)})^{-1}(y) = (\nabla \mathcal{F})^{-1}(y) = \zeta(y, v). \quad (3.7)$$
In particular, it follows from (3.5) that
$$\frac{\partial \psi_k}{\partial z_i}(y, v, v) = - \sum_{1 \leq j \leq r} \frac{\partial \psi_k}{\partial y_j}(y, v, v) \cdot \frac{\partial \xi_j}{\partial z_i}(\psi(y, v, v), v, v) \quad (3.8)$$
$$= - \sum_{1 \leq j \leq r} \frac{\partial \xi_k}{\partial y_j}(y, v) \cdot \frac{\partial^2 \mathcal{F}}{\partial z_i \partial u_j}(\zeta(y, v), v)$$
we obtain
\[ \zeta(\nabla F_z(u), z) = (\nabla F_z)^{-1}(\nabla F_z(u)) = u, \]
holds for all \( y \in B_r(y_0) \) and \( v \in B_{3\nu}(z_0) \). Then by taking the derivative with respect to \( z_i \) of the identity
\[
\zeta(\nabla F_z(u), z) = (\nabla F_z)^{-1}(\nabla F_z(u)) = u,
\]
we obtain
\[
0 = \sum_{1 \leq j \leq r} \frac{\partial \zeta_k}{\partial y_j}(\nabla F_z(u), z) \cdot \frac{\partial^2 F}{\partial z_i \partial u_j}(u, z) + \frac{\partial \zeta_k}{\partial z_i}(\nabla F_z(u), z)
\]
for each \( 1 \leq k \leq r \), which combined with (3.8) implies that
\[
\frac{\partial \zeta_k}{\partial z_i}(y, v) = - \sum_{1 \leq j \leq r} \frac{\partial \zeta_k}{\partial y_j}(y, v) \cdot \frac{\partial^2 F}{\partial z_i \partial u_j}(\zeta(y, v), v) = \frac{\partial \psi_k}{\partial z_i}(y, v, v) \tag{3.9}
\]
holds for all \( y \in B_r(y_0) \) and \( v \in B_{3\nu}(z_0) \). Therefore, it follows from (3.4), (3.6), (3.7) and (3.9) that
\[
\frac{\partial}{\partial z_i} \mathcal{L}_z^*(y, v)|_{z=v} = \frac{\partial}{\partial z_i} (y \cdot \psi(y, z, v) - \mathcal{L}(\psi(y, z, v), z, v))|_{z=v} = \sum_{1 \leq k \leq r} y_k \frac{\partial \psi_k}{\partial z_i}(y, v, v)
\]
\[
- \sum_{1 \leq k \leq r} \frac{\partial \mathcal{L}}{\partial u_k}(\psi(y, v, v), v, v) \frac{\partial \psi_k}{\partial z_i}(y, v, v) - \frac{\partial \mathcal{L}}{\partial z_i}(\psi(y, v, v), v, v)
\]
\[
= \sum_{1 \leq k \leq r} y_k \frac{\partial \zeta_k}{\partial z_i}(y, v) - \sum_{1 \leq k \leq r} \frac{\partial F}{\partial u_k}(\zeta(y, v), v) \frac{\partial \zeta_k}{\partial z_i}(y, v) - \frac{\partial F}{\partial z_i}(\zeta(y, v), v)
\]
\[
= \frac{\partial}{\partial z_i} (y \cdot \zeta(y, z) - F(\zeta(y, z), z))|_{z=v}
\]
\[
= \frac{\partial}{\partial z_i} \mathcal{F}_z^*(y)|_{z=v}
\]
holds for all \( y \in B_r(y_0) \) and \( v \in B_{3\nu}(z_0) \). Thus we have obtained statement (v), and statement (vi) follows by Taylor’s theorem. \( \square \)

3.2. Oscillatory integrals. The following is a non-stationary phase estimate \[\text{[8, Theorem 7.7.1].}\]

Lemma 3.4 (Non-stationary phase). Let \( m \geq 1, \ell \geq 2 \) and \( \mathcal{V} \subseteq \mathbb{R}^m \) a bounded open set. Let \( \omega \in C_0^{\ell-1}(\mathbb{R}^m) \) with \( \text{supp } \omega \subseteq \mathcal{V} \) and \( \xi \in C^\ell(\mathcal{V}) \) with
\[
\nabla \xi(t) \neq 0 \quad (t \in \text{supp } \omega).
\]
Let $\beta > 0$. Then
\[
\int_{\mathbb{R}^m} \omega(t)e(\beta \xi(t))dt \ll \beta^{-\ell+1},
\]
where the implicit constant depends only on $m$, $\ell$,
\[
\min_{t \in \text{supp } \omega} |\nabla \xi(t)|, \quad \|\xi\|_{C^\ell(\mathbb{R})} \quad \text{and} \quad \|\omega\|_{C^{\ell-1}(\mathbb{R}^m)}.
\]

We also need the following stationary phase estimate [8, Theorem 7.7.5].

Given an invertible symmetric matrix, by its signature we mean the number of positive eigenvalues minus the number of negative eigenvalues.

**Lemma 3.5** (Stationary phase). Let $1 \leq r \leq 2$, $\ell \geq 7$ and $\mathcal{U} \subseteq \mathbb{R}^r$ a bounded open set. Let $\omega \in C^{\ell-1}(\mathbb{R}^r)$ with $\text{supp } \omega \subseteq \mathcal{U}$ and $\varphi \in C^\ell(\mathcal{U})$. Suppose there exists $y_c \in \mathcal{U}$ satisfying $\det H_\varphi(y_c) \neq 0$, $\nabla \varphi(y_c) = 0$
\[
\nabla \varphi(y) \neq 0 \quad (y \in \text{supp } \omega \setminus \{y_c\}).
\]
Let $\lambda > 0$, $\sigma$ be the signature of $H_\varphi(y_c)$ and $\Delta = |\det H_\varphi(y_c)|$. Then
\[
\int_{\mathbb{R}^r} \omega(y)e(\lambda \varphi(y))dy = \Delta^{-\frac{\ell}{2}} \lambda^{-\frac{\ell}{2}} e \left( \lambda \varphi(y_c) + \frac{\sigma}{8} \right) (\omega(y_c) + O(\lambda^{-1})),
\]
where the implicit constant depends only on $\ell$, $\Delta$,
\[
\sup_{y \in \mathcal{U}} \frac{|y - y_c|}{|\nabla \varphi(y)|}, \quad \|\varphi\|_{C^\ell(\mathcal{U})} \quad \text{and} \quad \|\omega\|_{C^{\ell-1}(\mathbb{R}^r)}.
\]

3.3. Rational points near curves in $\mathbb{R}^2$/surfaces in $\mathbb{R}^3$.

**Proposition 3.6.** Let $1 \leq r \leq 2$, $d > r$, $\ell \geq 7$, $\mu, \nu > 0$ sufficiently small, $u_0 \in \mathbb{R}^r$, $0 \in \mathbb{R}^{d-r}$ and $G_\nu \in C^\ell(B^{\mu}(u_0))$ for each $v \in B_{3\nu}(0)$. Let $\mathcal{W} \in C^\infty(\mathbb{R}^r)$ such that
\[
\text{supp } \mathcal{W} \subseteq B^{\mu}(u_0)
\]
and there exists $\varrho_1 > 0$ satisfying
\[
\text{dist}(\partial B^{\mu}(u_0), \partial(\text{supp } \mathcal{W})) \geq \varrho_1.
\]
Suppose $\nabla G_\nu$ is a $C^{\ell-1}$-diffeomorphism on $B^{\mu}(u_0)$ for each $v \in B_{3\nu}(0)$,
\[
\sup_{v \in B_{3\nu}(0)} \|G_\nu\|_{C^\ell(B^{\mu}(u_0))} \ll 1,
\]
and there exist $c_1, c_2, \varrho_2 > 0$ satisfying
\[
c_1 \leq |\det H_{G_\nu}(u)| \leq c_2 \quad (u \in B^{\mu}(u_0), v \in B_{3\nu}(0))
\]
and
\[
\inf_{v \in B_{3\nu}(0)} \text{dist}(\partial(\nabla G_\nu(B^{\mu}(u_0))), \partial(\nabla G_\nu(\text{supp } \mathcal{W}))) \geq \varrho_2.
\]
Let $P, T \geq 1$. Then

$$\sup_{v \in B_{3\nu}(0)} \sum_{\frac{T}{2} \leq j \leq T} \sum_{h \in \mathbb{Z}} \mathcal{W} \left( \frac{h}{j} \right) \min(P, \| jG_v(h/j) \|^{-1}) \ll T^{r+1} P^\varepsilon + P T^{r+\varepsilon}$$

for any $\varepsilon > 0$.

**Proof.** Let $X \geq 2$ and $v \in B_{3\nu}(0)$. If $r = 1$, then [[10]] Theorems 1 and 2 imply

$$\sum_{\frac{T}{2} \leq j \leq T} \sum_{h \in \mathbb{Z}} \mathcal{W} \left( \frac{h}{j} \right) \ll X^{-1+\varepsilon} T^2 + T$$

for any $\varepsilon > 0$, where the implicit constant is independent of $v$. Therefore, it follows that

$$\sum_{\frac{T}{2} \leq j \leq T} \sum_{h \in \mathbb{Z}} \mathcal{W} \left( \frac{h}{j} \right) \ll \sum_{1 \leq t \leq T} \sum_{\frac{T}{2} \leq j \leq T} \sum_{m, h \in \mathbb{Z} \atop \gcd(m, h, j) = 1} \mathcal{W} \left( \frac{h}{j} \right)$$

$$= \sum_{1 \leq t \leq T} \sum_{\frac{T}{2} \leq j \leq T} \sum_{m, h \in \mathbb{Z} \atop \gcd(m, h, j) = t} \mathcal{W} \left( \frac{h}{j} \right)$$

$$\ll \sum_{1 \leq t \leq T} (Xt)^{-1+\varepsilon} \left( \frac{T}{t} \right)^2 + T$$

$$\ll X^{-1+\varepsilon} T^2 + T \log T.$$

If $r = 2$, then [[9], Theorem 1] implies

$$\sum_{\frac{T}{2} \leq j \leq T} \sum_{h \in \mathbb{Z}^2 \atop \| jG_v(h/j) \| \leq X^{-1}} \mathcal{W} \left( \frac{h}{j} \right) \ll X^{-1} T^3 + T^{2+\varepsilon} \quad (3.11)$$

for any $\varepsilon > 0$, where the implicit constant depends only on $\varepsilon, u_0, \mu, \nu, c_1, c_2, \varrho_1, \varrho_2, \| G_v \|_{C^r(B_{u_0}(u_0))}$ and $\| \mathcal{W} \|_{C^{r-1}(\mathbb{R}^r)}$. Thus by combining (3.10) and (3.11) we obtain

$$\sup_{v \in B_{3\nu}(0)} \sum_{\frac{T}{2} \leq j \leq T} \sum_{h \in \mathbb{Z}^2 \atop \| jG_v(h/j) \| \leq X^{-1}} \mathcal{W} \left( \frac{h}{j} \right) \ll X^{-1+\varepsilon} T^{r+1} + T^{r+\varepsilon} \quad (3.12)$$
Therefore, it follows that
\[
\sum_{\frac{T}{2} \leq j \leq T} \sum_{h \in \mathbb{Z}^r} \mathbb{W} \left( \frac{h}{j} \right) \min \left( P, \|jG_v(h/j)\|^{-1} \right) \tag{3.13}
\]
\[
\ll P \left( P^{-1+\epsilon} T^{r+1} + T^{r+\epsilon} \right) + \sum_{\frac{T}{2} \leq j \leq T} \sum_{h \in \mathbb{Z}^r} \mathbb{W} \left( \frac{h}{j} \right) \|jG_v(h/j)\|^{-1}.
\]

In order to estimate the final term, we split the interval \([P^{-1}, 1/2]\) into dyadic intervals. Since the sum is an empty sum if \(P^{-1} > 1/2\), we assume \(P^{-1} \leq 1/2\). Then by (3.12) we have
\[
\sum_{\frac{T}{2} \leq j \leq T} \sum_{h \in \mathbb{Z}^r} \mathbb{W} \left( \frac{h}{j} \right) \|jG_v(h/j)\|^{-1} \ll \sum_{0 \leq i \leq \log P \log 2} \left( \frac{P}{2^i} \right) T^{r+1} \left( P^{1-\epsilon} T^{r+1} + T^{r+\epsilon} \right)
\]
\[
\ll T^{r+1} (\log P) P^\epsilon + PT^{r+\epsilon}.
\]

Therefore, by substituting this estimate into (3.13) we obtain
\[
\sup_{v \in B_3(0)} \sum_{\frac{T}{2} \leq j \leq T} \sum_{h \in \mathbb{Z}^r} \mathbb{W} \left( \frac{h}{j} \right) \min \left( P, \|jG_v(h/j)\|^{-1} \right) \ll T^{r+1} P^\epsilon + PT^{r+\epsilon}.
\]

\[\square\]

4. Proof of Theorem 1.8

Let \(d \in \mathbb{N}, \ell \geq \max(d + 1, 7), \mathcal{O} \subseteq \mathbb{R}^d\) an open set and \(f \in C^{\ell+2}(\mathcal{O})\). Let \(x_0 \in \mathcal{O}\) and suppose \(H_f(x_0)\) is a non-zero matrix. We have the following trivial lemma regarding symmetric matrices.

Lemma 4.1. Suppose \(A = [a_{i,j}]_{1 \leq i, j \leq d}\) is a non-zero \(d \times d\) real symmetric matrix. Then there exist \(1 \leq i \leq d\) such that \(a_{i,i} \neq 0\) or \(1 \leq j < k \leq d\) such that \(a_{j,j} = a_{k,k} = 0\) and \(a_{j,k} \neq 0\).

Using this lemma with \(H_f(x_0)\), we denote \(x = (y, z)\) so that \(x_0 = (y_0, z_0)\) and \(H_{f_{x_0}}(y_0)\) is an invertible \(r \times r\) matrix, where \(r = 1\) in the former case and \(r = 2\) in the latter case. Since the statement of Theorem 1.8 when
(d, r) = (1, 1) and (2, 2) follow from [13, Theorem 3] and [9, Theorem 1] respectively, we assume from this point on that

\[ 1 \leq r \leq 2 \quad \text{and} \quad 1 \leq r < d. \]

We begin by showing that we may assume without of loss generality \( z_0 = 0 \). Let \( \nu > 0 \) be sufficiently small. Let \( a \in \mathbb{Z}^{d-r} \) and \( b \in \mathbb{N} \) satisfying

\[ \left| z_0 - \frac{a}{b} \right| < \nu. \]

Let us define \( \tilde{f}(y, z) = f(y, z + \frac{a}{b}) \).

Then

\[
N_f(Q, \delta) = \sum_{1 \leq q \leq Q} \sum_{(h, k) \in \mathbb{Z}^d} 1_{B_{2\nu}(y_0)} \left( \frac{h}{q} \right) 1_{B_{2\nu}(z_0)} \left( \frac{k}{q} \right) \leq \sum_{1 \leq q \leq Q} \sum_{(h, k) \in \mathbb{Z}^d} 1_{B_{2\nu}(y_0)} \left( \frac{h}{q} \right) 1_{B_{2\nu}(a/b)} \left( \frac{k}{q} \right) \leq \sum_{1 \leq q \leq bQ} \sum_{(h, k) \in \mathbb{Z}^d} 1_{B_{2\nu}(y_0)} \left( \frac{h}{q} \right) 1_{B_{2\nu}(0)} \left( \frac{k}{q} \right),
\]

where the final inequality follows on noting that

\[
\left\| q\tilde{f} \left( \frac{h}{q}, \frac{k}{q} - \frac{a}{b} \right) \right\| \leq \delta \quad \text{and} \quad \frac{k}{q} \in B_{2\nu} \left( \frac{a}{b} \right)
\]

imply

\[
\left\| bq\tilde{f} \left( \frac{bh}{bq}, \frac{bk - qa}{bq} \right) \right\| \leq b\delta \quad \text{and} \quad \frac{bk - qa}{bq} \in B_{2\nu} (0).
\]

Since \( \det H_{f_{z_0}}(y_0) \neq 0 \), by continuity of the entries of \( H_{f_z}(y) \) it follows that \( \det H_{f_0}(y_0) = \det H_{f_{z_0}}(y_0) \neq 0 \) provided \( \nu \) is sufficiently small. Therefore, in order to establish Theorem 1.8 it suffices to prove the result under the additional hypothesis \( z_0 = 0 \); we assume \( z_0 = 0 \) from here on.
Let \( w \in C_0^\infty(\mathbb{R}^r) \) and \( \chi \in C_0^\infty(\mathbb{R}^{d-r}) \) such that \( \text{supp } w \subseteq B_{2\nu}(y_0) \), \( \text{supp } \chi \subseteq B_{2\nu}(0) \),

\[
1_{B_r(y_0)}(y) \leq w(y) \quad (y \in \mathbb{R}^r) \quad \text{and} \quad 1_{B_r(0)}(z) \leq \chi(z) \quad (z \in \mathbb{R}^{d-r}).
\]

For any \( P \geq 1 \), let us define

\[
\mathcal{N}(P, \delta) = \sum_{\frac{P}{2} \leq q \leq P} \sum_{(h,k) \in \mathbb{Z}^d} \sum_{\|qf(h/k,q/k)\| \leq \delta} w\left(\frac{h}{q}\right) \chi\left(\frac{k}{q}\right).
\]

It is clear that

\[
\sum_{\frac{P}{2} \leq q \leq P} \sum_{(h,k) \in \mathbb{Z}^d} 1_{B_r(y_0)}\left(\frac{h}{q}\right) 1_{B_r(0)}\left(\frac{k}{q}\right) \leq \mathcal{N}(P, \delta). \tag{4.2}
\]

Next we apply Lemma 3.3 with

\[
\mathcal{F}(u,z) = f_z^*(u) \quad \text{and} \quad u_0 = \nabla f_0(y_0).
\]

Since

\[
\det H_{\mathcal{F}_0}(u_0) = \det H_{f_0}(u_0) = \frac{1}{\det H_{f_0}(y_0)} \neq 0, \tag{4.3}
\]

it follows by Lemma 3.2 that there exists \( \mu > 0 \) such that \( \mathcal{F} \in C^{\mu+1}(B_\mu(u_0) \times B_\mu(0)) \); therefore, the hypotheses to apply Lemma 3.3 are satisfied. Let \( \mathcal{L} \) and \( g \) be as in the statement of Lemma 3.3. It is easy to see that \( \mathcal{L} \in C^\mu(B_\mu(u_0) \times B_\mu(0) \times B_\mu(0)) \) and

\[
\det H_{\mathcal{L}_0}(u_0) = \det H_{\mathcal{F}_0}(u_0) \neq 0.
\]

For \( \mu \) sufficiently small, it further follows by Lemma 3.2 that

\[
g_{z,v}^*(u) = \mathcal{L}_{(z,v)}(u) = \mathcal{F}_{(z,v)}(u) + \sum_{1 \leq i \leq d-r} \frac{\partial \mathcal{F}}{\partial z_i}(u,v) z_i \tag{4.4}
\]

holds for all \( u \in B_\mu(u_0), z \in B_\mu(0) \) and \( v \in B_\mu(0) \), where

\[
\mathcal{F}_{(z,v)}(u) = \mathcal{F}(u,v) - \sum_{1 \leq i \leq d-r} \frac{\partial \mathcal{F}}{\partial z_i}(u,v) v_i. \tag{4.5}
\]

Let \( \tau > 0 \) be sufficiently small and \( \nu \) sufficiently small with respect to \( \tau \). Then by Lemma 3.3 we have \( g \in C^{\mu+1}(B_\tau(y_0) \times B_{2\nu}(0) \times B_{3\nu}(0)) \), \( g_{(z,v)} \in C^\mu(B_\tau(y_0)) \) and \( \nabla g_{(z,v)} \) is a \( C^{\mu+1} \)-diffeomorphism on \( B_\tau(y_0) \) for each \( z \in B_{2\nu}(0) \) and \( v \in B_{3\nu}(0) \), there exist \( c_- \), \( c_+ > 0 \) such that

\[
c_\tau \leq |\det H_{g_{(z,v)}}(y)| \leq c_+ \quad (y \in B_\tau(y_0), z \in B_{2\nu}(0), v \in B_{3\nu}(0)). \tag{4.6}
\]

and there exists \( \varrho > 0 \) such that

\[
\text{dist}(\partial(\nabla g_{(z,v)}(B_\tau(y_0))), \partial(\nabla g_{(z,v)}(B_{2\nu}(y_0)))) \geq \varrho \quad (z \in B_{2\nu}(0), v \in B_{3\nu}(0)), \tag{4.7}
\]
Furthermore, there exists $C > 0$ such that
\begin{equation}
|\mathcal{F}_z^*(y) - \mathcal{L}^*_{(z,v)}(y)| = |f(y,z) - g(y,z,v)| \leq C|z - v|^2
\end{equation}
holds for all $y \in B_r(y_0)$, $z \in B_{2\nu}(0)$ and $v \in B_{3\nu}(0)$.

Let
\begin{equation}
\mathcal{X} = \bigcup_{z \in B_{2\nu}(0)} \nabla g_{(z,v)}(B_{2\nu}(y_0)).
\end{equation}
We may further assume without loss of generality that
\begin{equation}
\mathcal{X} \subseteq \bigcap_{z \in B_{2\nu}(0)} \nabla g_{(z,v)}(B_r(y_0)),
\end{equation}
and by Lemma 3.1 if we denote $\psi(u,z,v) = (\nabla g_{(z,v)})^{-1}(u)$, then
\begin{equation}
\psi_k \in C^{\ell-2}(\mathcal{X} \times B_{2\nu}(0) \times B_{3\nu}(0))
\end{equation}
for each $1 \leq k \leq r$. We record the following simple statements regarding $\mathcal{X}$ and $G^*_v$.

**Lemma 4.2.** Let $\nu$ be sufficiently small with respect to $\mu$. Then $\mathcal{X}$ is contained in the domain of $G^*_v$ for all $v \in B_{3\nu}(0)$, i.e.
\begin{equation}
\mathcal{X} \subseteq B_\mu(u_0).
\end{equation}
Furthermore, there exists $\varrho'_1 > 0$ such that
\begin{equation}
\text{dist}(\partial B_\mu(u_0), \partial \mathcal{X}) \geq \varrho'_1.
\end{equation}

**Proof.** The result follows easily from the fact that
\begin{align*}
\nabla g_{(0,0)}(y_0) &= \nabla \mathcal{F}^*_{(0,0)}(y_0) \\
&= (\nabla \mathcal{L}_{(0,0)})^{-1}(y_0) \\
&= (\nabla \mathcal{F}_0)^{-1}(y_0) \\
&= (\nabla f_0^* )^{-1}(y_0) \\
&= \nabla f_0(y_0) \\
&= u_0.
\end{align*}

\hfill \Box

**Lemma 4.3.** Let $\mu$ be sufficiently small and $\nu$ sufficiently small with respect to $\mu$. Then we have $\nabla G^*_v$ is a $C^{\ell-1}$-diffeomorphism on $B_\mu(u_0)$ for each $v \in B_{3\nu}(0)$,
\begin{equation}
\sup_{v \in B_{3\nu}(0)} \|\nabla G^*_v\|_{C^{\ell}(B_\mu(u_0))} \ll 1,
\end{equation}
and there exist $c_1, c_2, \varrho'_2 > 0$ satisfying
\[ c_1 \leq |\det H_{\varphi'_2}(\mathbf{u})| \leq c_2 \quad (\mathbf{u} \in B_{\mu}(\mathbf{u}_0), \mathbf{v} \in B_{3\nu}(0)), \]
and
\[ \inf_{\mathbf{v} \in B_{3\nu}(0)} \text{dist}(\partial(\nabla \varphi'_2(B_{\mu}(\mathbf{u}_0))), \partial(\nabla \varphi'_2(\mathbb{R}^n))) \geq \varrho'_2. \]

Proof. Let $\mathbf{v} \in B_{3\nu}(0)$. Since we have (4.3) and
\[ \left| \frac{\partial^2 \varphi'_2}{\partial u_j \partial u_k}(\mathbf{u}) - \frac{\partial^2 \varphi}{\partial u_j \partial u_k}(\mathbf{u}, \mathbf{v}) \right| \leq 3\nu \sum_{1 \leq i \leq d-r} \left| \frac{\partial^1 \varphi}{\partial u_j \partial u_k \partial z_i}(\mathbf{u}, \mathbf{v}) \right| \quad (\mathbf{u} \in B_{\mu}(\mathbf{u}_0)) \]
for each $1 \leq j, k \leq r$, it follows that there exist $c_1, c_2 > 0$ satisfying
\[ c_1 \leq |\det H_{\varphi'_2}(\mathbf{u})| \leq c_2 \quad (\mathbf{u} \in B_{\mu}(\mathbf{u}_0), \mathbf{v} \in B_{3\nu}(0)) \]
provided $\mu$ and $\nu$ are sufficiently small. In particular, we have $\det H_{\varphi'_2}(0) \neq 0.$

If we denote $G(\mathbf{u}, \mathbf{v}) = \varphi'_2(\mathbf{u})$, then it easily follows from the definition of $\varphi'_2$ and the fact that $\varphi \in C^{\ell+1}(B_{\mu}(\mathbf{u}_0) \times B_{\mu}(0))$ that $G \in C^{\ell}(B_{\mu}(\mathbf{u}_0) \times B_{\mu}(0))$. Therefore, by Lemma 3.1 it follows that there exists $\kappa > 0$ such that $\nabla \varphi'_2$ is a $C^{\ell-1}$-diffeomorphism on $B_{\kappa}(\mathbf{u}_0)$ for each $\mathbf{v} \in B_{\kappa}(0)$, and the first statement follows provided $\mu < \kappa$ and $\nu < \kappa/3$. The second statement follows easily from the first, and the final statement from the first and Lemma 4.2. \hfill \square

Let $0 < \eta \leq \nu$ be a parameter to be set in due course. We introduce the following smooth weight function
\[ \chi_{\eta}(\mathbf{z}, \mathbf{v}) = \prod_{1 \leq i \leq d-r} \omega_0(z_i) \cdot \chi(\mathbf{v} + \eta \mathbf{z}), \]
where
\[ \omega_0(t) = \begin{cases} \left( \int_{-1}^1 e^{-(1-s^2)^{-1}} \, ds \right)^{-1} e^{-(1-t^2)^{-1}} & \text{if } |t| < 1, \\ 0 & \text{if } |t| \geq 1. \end{cases} \]

We have
\[ \int_{\mathbb{R}} \omega_0(-t) \, dt = \int_{\mathbb{R}} \omega_0(t) \, dt = 1 \]
and
\[ \omega'_0(t) = \frac{d\omega_0}{dt}(t) \quad (4.12) \]
\[ = \begin{cases} \left( \int_{-1}^1 e^{-(1-s^2)^{-1}} \, ds \right)^{-1} (1 - t^2)^{-2}(-2t)e^{-(1-t^2)^{-1}} & \text{if } |t| < 1, \\ 0 & \text{if } |t| \geq 1. \end{cases} \]
It is clear that \( \omega'_0(t) > 0 \) if \(-1 < t < 0, \omega'_0(t) < 0 \) if \(0 < t < 1 \) and \( \omega'_0(t) = 0 \) otherwise. Also it is proved in [6, Lemma 2] that
\[
\chi(z) = \eta^{-d+r} \int_{\mathbb{R}^{d-r}} \chi_{\eta^{-1}(z-v)}(v)dv
\]
\[
= \eta^{-d+r} \int_{|z-v| \leq \eta} \chi_{\eta^{-1}(z-v)}(v)dv
\]
\[
= \eta^{-d+r} \int_{B_{3\nu}(0)} \chi_{\eta^{-1}(z-v)}(v)dv.
\]
In particular, we only need to consider
\[
v \in B_{3\nu}(0).
\]
Then
\[
\mathcal{N}(P, \delta) = \eta^{-d+r} \int_{\mathbb{R}^{d-r}} \sum_{\frac{P}{2} \leq q \leq P} \sum_{\|qf(h/q, k/q)\| \leq \delta} w\left(\frac{h}{q}\right) \chi_{\eta^{-1}\left(\frac{k}{q} - v\right)}(v)dv.
\]
Given \( P/2 \leq q \leq P, v \in B_{3\nu}(0) \) and \( h/q \in \text{supp } w \), if \( k \in \mathbb{Z}^{d-r} \) satisfies
\[
\chi_{\eta^{-1}\left(\frac{k}{q} - v\right)}(v) \neq 0,
\]
then
\[
\frac{k}{q} \in \text{supp } \chi \subseteq B_{2\nu}(0) \quad \text{and} \quad \left|\frac{k}{q} - v\right| \leq \eta,
\]
and by (4.13) we have
\[
\|qf(h/q, k/q) - qg(h/q, k/q, v)\| \leq Cq\eta^2 \leq CP\eta^2.
\]
It follows that
\[
\|qf(h/q, k/q)\| \leq \delta \quad \text{implies} \quad \|qg(h/q, k/q, v)\| \leq \delta + CP\eta^2.
\]
Therefore, we obtain
\[
\mathcal{N}(P, \delta) \leq \mathcal{M}(P, \delta + C\eta^2),
\]
where
\[
\mathcal{M}(P, \alpha) = \eta^{-d+r} \int_{\mathbb{R}^{d-r}} \sum_{\frac{P}{2} \leq q \leq P} \sum_{\|qf(h/q, k/q, v)\| \leq \alpha} w\left(\frac{h}{q}\right) \chi_{\eta^{-1}\left(\frac{k}{q} - v\right)}(v)dv.
\]
We present the proof of the following proposition in the next section.

**Proposition 4.4.** Let \( 0 < \alpha \leq 1/2 \). Then
\[
\mathcal{M}(P, \alpha) \ll \alpha P^{d-1} + \eta^{-d} + \left(P^{d-\frac{d}{2}} \alpha^{-\frac{d}{2}} + P^{d+1-\frac{d}{2}} \alpha^{1-\frac{d}{2}}\right) P^{\epsilon \alpha^{-\epsilon}}
\]
for any \( \epsilon > 0 \).
In order to prove Theorem 1.8, we combine (4.2), (4.14) and this proposition with
\[ \alpha = \delta + CP\eta^2 \quad \text{and} \quad \eta = \nu / P. \]
In particular, it follows that
\[ \mathcal{M} \left( P, \delta + \frac{C\nu^2}{P} \right) \ll \delta P^d + 1 + \delta^{1-\frac{r}{2}} P^{d+1-\frac{r}{2}+\epsilon} + P^{d+\epsilon} \]
provided
\[ \delta + \frac{C\nu^2}{P} \leq \frac{1}{2}. \]
We note that this bound holds even if
\[ \delta + \frac{C\nu^2}{P} > \frac{1}{2}, \]
since
\[ \mathcal{M} \left( P, \delta + \frac{C\nu^2}{P} \right) = \mathcal{M} \left( P, \frac{1}{2} \right) \ll P^{d+1} \ll \left( \delta + \frac{C\nu^2}{P} \right) P^{d+1} \ll \delta P^{d+1} + P^d \]
in this case. As a result, we obtain
\begin{align*}
N_f(Q, \delta) &\ll \sum_{0 \leq s \leq \log Q} \sum_{\frac{Q}{2^{s+1}} \leq q \leq \frac{Q}{2}} \sum_{(h,k) \in \mathbb{Z}^d \atop \|f(h/q,k/q)\| \leq \delta} \mathbb{1}_{B_r(y_0)} \left( \frac{h}{q} \right) \mathbb{1}_{B_r(0)} \left( \frac{k}{q} \right) \\
&\ll \sum_{0 \leq s \leq \log^2 Q} \mathcal{N} \left( \frac{Q}{2s}, \delta \right) \\
&\ll \sum_{0 \leq s \leq \log^2 Q} \mathcal{M} \left( \frac{Q}{2s}, \delta + \frac{C\nu^2}{P} \right) \\
&\ll \delta \sum_{0 \leq s \leq \log^2 Q} \left( \frac{Q}{2s} \right)^{d+1} + \delta^{1-\frac{r}{2}} \sum_{0 \leq s \leq \log^2 Q} \left( \frac{Q}{2s} \right)^{d+1-\frac{r}{2}+\epsilon} + Q^{d+\epsilon} \\
&\ll \delta Q^{d+1} + \delta^{1-\frac{r}{2}} Q^{d+1-\frac{r}{2}+\epsilon} + Q^{d+\epsilon}.
\end{align*}
Let \( \vartheta = 2\epsilon / r. \) If \( 0 \leq \delta \leq Q^{-1+\vartheta} \), then
\[ N_f(Q, \delta) \ll N_f(Q, Q^{-1+\vartheta}) \ll Q^{d+\vartheta}. \]
If \( Q^{-1+\vartheta} < \delta \leq 1/2 \), then
\begin{align*}
N_f(Q, \delta) &\ll \delta Q^{d+1} \left( 1 + \delta^{-\frac{r}{2}} Q^{-\frac{r}{2}+\epsilon} \right) + Q^{d+\epsilon} \\
&\ll \delta Q^{d+1} \left( 1 + Q^{\frac{\vartheta}{2}+\epsilon} \right) + Q^{d+\epsilon} \\
&\ll \delta Q^{d+1} + Q^{d+\vartheta}. 
\end{align*}
Therefore, it follows that
\[ N_f(Q, \delta) \ll \delta Q^{d+1} + Q^{d+\varepsilon} \]
holds for all \( 0 \leq \delta \leq 1/2 \); this completes the proof of Theorem 1.8.

5. Proof of Proposition 4.4

In this section, for simplicity we denote
\[ g(y, z) = g(y, z, v) \quad \text{and} \quad g_z(y) = g_{(z, v)}(y). \]
Let \( 0 < \alpha \leq 1/2 \) and
\[ 1_{[-\alpha, \alpha] + Z}(\theta) = \begin{cases} 1 & \text{if } \|\theta\| \leq \alpha, \\ 0 & \text{otherwise}. \end{cases} \]
Let
\[ J = \left\lfloor \frac{1}{2\alpha} \right\rfloor \]
and \( \mathcal{F}_J \) denote the Fejér kernel defined by
\[ \mathcal{F}_J(\theta) = J^{-2} \left| \sum_{-J \leq j \leq J} e(j\theta) \right|^2 = \left( \frac{\sin(\pi J \theta)}{J \sin(\pi \theta)} \right)^2 = \sum_{-J \leq j \leq J} \frac{J - |j|}{J^2} e(j\theta) \quad (\theta \in \mathbb{R}). \]
Since
\[ \left( \frac{\sin(\pi J \theta)}{J \sin(\pi \theta)} \right)^2 \geq \left( \frac{2\pi^{-1}}{J \pi \|\theta\|} \right)^2 \geq \frac{4}{\pi^2} \quad (0 < \|\theta\| \leq \alpha), \]
it follows that
\[ 1_{[-\alpha, \alpha] + Z}(\theta) \leq \frac{\pi^2}{4} \mathcal{F}_J(\theta) \quad (\theta \in \mathbb{R}). \quad (5.1) \]
It is clear by (4.13) that
\[
\eta^{-d+r} \int_{\mathbb{R}^{d-r}} \sum_{\frac{h}{q} \leq P} \sum_{(h, k) \in \mathbb{Z}^d} w\left( \frac{h}{q} \right) \chi_q \left( \eta^{-1} \left( \frac{k}{q} - v \right), v \right) dv
\begin{align*}
&= \sum_{\frac{h}{q} \leq P} \sum_{(h, k) \in \mathbb{Z}^d} w\left( \frac{h}{q} \right) \chi\left( \frac{k}{q} \right) \\
&\ll P^{d+1}.
\end{align*}
Therefore, by (5.1) and this inequality we have
\[ \mathcal{M}(P, \alpha) \ll \frac{P^{d+1}}{J} \] (5.2)

\[ + \frac{\eta^{-d+r}}{J} \sum_{1 \leq j < J} \left| \int_{\mathbb{R}^{d-r}} \sum_{P \leq q \leq P} \sum_{\frac{h}{q}, \frac{k}{q} \in \mathbb{Z}} w\left(\frac{h}{q}\right) \chi_{\eta} \left(\eta^{-1} \left(\frac{k}{q} - v\right), v\right) e\left(qjg\left(\frac{h}{q} \cdot \frac{k}{q}\right)\right) dv \right| . \]

By the \((d - r)\)-dimensional Poisson summation formula, we obtain
\[ \eta^{-d+r} \sum_{(h,k) \in \mathbb{Z}^d} w\left(\frac{h}{q}\right) \chi_{\eta} \left(\eta^{-1} \left(\frac{z}{q} - v\right), v\right) e\left(qjg\left(\frac{h}{q} \cdot \frac{z}{q}\right) - k \cdot z\right) dz \]
\[ = \eta^{-d+r} \sum_{(h,k) \in \mathbb{Z}^d} \int_{\mathbb{R}^{d-r}} w\left(\frac{h}{q}\right) \chi_{\eta} \left(\eta^{-1} (z - v), v\right) e\left(qjg\left(\frac{h}{q} \cdot \frac{z}{q}\right) - qk \cdot z\right) dz \]
\[ = q^{d-r} \sum_{(h,k) \in \mathbb{Z}^d} \int_{\mathbb{R}^{d-r}} w\left(\frac{h}{q}\right) \chi_{\eta} (z - \eta^{-1} v, v) e\left(qjg\left(\frac{h}{q} \eta z\right) - \eta q k \cdot z\right) dz. \]

Then (5.2) becomes
\[ \mathcal{M}(P, \alpha) \ll \frac{P^{d+1}}{J} \] (5.4)

\[ + \frac{1}{J} \sum_{1 \leq j < J} \left| \int_{\mathbb{R}^{d-r}} \sum_{\frac{h}{q}, \frac{k}{q} \in \mathbb{Z}} q^{d-r} \sum_{(h,k) \in \mathbb{Z}^d} \int_{\mathbb{R}^{d-r}} w\left(\frac{h}{q}\right) \chi_{\eta} (z - \eta^{-1} v, v) e\left(qjg\left(\frac{h}{q} \eta z\right) - \eta q k \cdot z\right) dz dv \right| . \]

Let \(m \in \mathbb{R}_{\geq 1}\) be a constant satisfying
\[ \sup_{y \in \text{supp} \ w} \frac{1}{z \in B_2(0)} \left| \nabla g_y(z) \right| \leq m, \]
where \(\nu\) is as in Section 4 and
\[ \nabla g_y(z) = \left( \frac{\partial g}{\partial z_1}(y, z), \ldots, \frac{\partial g}{\partial z_{d-r}}(y, z) \right). \]
It is clear from the definition of $\chi_\eta$ that

$$\chi_\eta(z - \eta^{-1}v, v) \neq 0$$

implies

$$\eta z \in \text{supp} \chi \subseteq B_{2\nu}(0) \quad \text{and} \quad v \in B_{3\nu}(0).$$

Given any $|k| > 2mJ$, we have that

$$\left| \nabla \left( \frac{j}{|k|} g_y(\eta z) - \frac{k}{|k|} \cdot z \right) \right| = \left| \frac{j}{|k|} \nabla g_y(\eta z) - \frac{k}{|k|} \right| \geq 1 - \frac{j |\nabla g_y(\eta z)|}{|k|} > \frac{1}{2}$$

holds for all $y \in \text{supp} w$, $\eta z \in B_{2\nu}(0)$ and $v \in B_{3\nu}(0)$. Therefore, for $|k| > 2m(P^2 + J)$ it follows from Lemma 3.4 (with $\ell = d + 1$) that

$$\left| \int_{\mathbb{R}^{d-r}} \chi_\eta(z - \eta^{-1}v, v) e \left( q g_y \left( \frac{h}{q}, \eta z \right) - \eta q k \cdot z \right) \, dz \right|$$

$$= \left| \int_{\mathbb{R}^{d-r}} \chi_\eta(z - \eta^{-1}v, v) e \left( \eta q |k| \left( \frac{j}{|k|} g_y \left( \frac{h}{q}, \eta z \right) - \frac{k}{|k|} \cdot z \right) \right) \, dz \right|$$

$$\ll \frac{1}{(\eta q |k|)^d}.$$

We use this estimate to deduce

$$\frac{1}{J} \sum_{1 \leq j \leq J} \int_{\mathbb{R}^{d-r}} \sum_{\frac{p}{q} \leq q \leq P} q^{d-r} \sum_{h \in \mathbb{Z}} w \left( \frac{h}{q} \right)$$

$$\sum_{k \in \mathbb{Z}^{d-r}} \left| \int_{\mathbb{R}^{d-r}} \chi_\eta(z - \eta^{-1}v, v) e \left( q g_y \left( \frac{h}{q}, \eta z \right) - \eta q k \cdot z \right) \, dz \right| \, dv$$

$$\ll \frac{1}{J} \sum_{1 \leq j \leq J} \sum_{\frac{p}{q} \leq q \leq P} q^d \sum_{k \in \mathbb{Z}^{d-r}} \frac{1}{(\eta q |k|)^d} \int_{B_{3\nu}(0)} 1 \, dv$$

$$\ll \eta^{-d} P \sum_{k \in \mathbb{Z}^{d-r}} \frac{1}{|k|^d} \sum_{|k| > 2m(P^2 + J)}$$

$$\ll \eta^{-d} P \sum_{k > 2m(P^2 + J)} k^{-d + d - r - 1}$$

$$\ll \eta^{-d} \sum_{k > 2m(P^2 + J)} k^{-r - \frac{1}{2}}$$

$$\ll \eta^{-d}.$$
By the $r$-dimensional Poisson summation formula, we obtain
\[
\sum_{h \in \mathbb{Z}^r} w\left(\frac{h}{q}\right) e\left(qjg\left(\frac{h}{q}, \eta z\right)\right) = \int_{\mathbb{R}^r} \sum_{h \in \mathbb{Z}^r} w\left(\frac{y}{q}\right) e\left(qjg\left(\frac{y}{q}, \eta z\right) - h \cdot y\right) dy
\]
\[
= q^r \int_{\mathbb{R}^r} \sum_{h \in \mathbb{Z}^r} w(y)e(qjg(y, \eta z) - qh \cdot y)dy.
\]

Therefore, by combining (5.4), (5.5) and this equality it follows that
\[
\mathcal{M}(P, \alpha) \ll \frac{P^{d+1}}{J} + \eta^{-d}
\]
\[
+ \frac{1}{J} \sum_{1 \leq j \leq J} \left| \int_{\mathbb{R}^{d-r}} \int_{\mathbb{R}^{d-r}} \sum_{p \leq q \leq P} q^d \chi_p(z - \eta^{-1}v, v) \Phi(\eta qz) \sum_{h \in \mathbb{Z}^r} I_z^{(v)}(q; j; h) dz dv \right|,
\]
where
\[
I_z^{(v)}(q; j; h) = \int_{\mathbb{R}^r} w(y)e(qjg(y, \eta z) - qh \cdot y)dy
\]
and
\[
\Phi(z) = \sum_{k \in [-2mP^2+J,2mP^2+J]^{d-r}} e(-k \cdot z).
\]

The following estimate becomes useful in dealing with $\Phi$.

**Lemma 5.1.** Let $P, J \geq 1$. Then
\[
\int_{[-P, P]^{d-r}} |\Phi(z)| dz \ll P^{d-r}(\log(P^2 + J))^{d-r}.
\]

**Proof.** The result is an immediate consequence of the following estimate
\[
\int_{-P}^P \left| \sum_{k \in [-2mP^2+J,2mP^2+J]} e(-kz) \right| dz \ll P \int_0^{1/2} \min(P^2 + J, \|z\|^{-1}) dz
\]
\[
\ll P \left( \int_0^{1/(P^2 + J)} (P^2 + J) dz + \int_{1/(P^2 + J)}^{1/2} \frac{1}{z} dz \right)
\]
\[
\ll P \log(P^2 + J).
\]
Let
\[ U = \{ y \in \mathbb{R}^r : w(y) \neq 0 \} \subseteq \text{supp } w \subseteq B_{2\nu}(y_0) \subseteq B_{\tau}(y_0) \]
and
\[ 0 < \rho < \varrho, \]
where \( \tau \) is as in Section 4 and \( \varrho \) is from (4.7). For each \( z \in \eta^{-1}B_{2\nu}(0) \) and \( v \in B_{3\nu}(0) \), let us define
\[ V_z^{(v)} = \nabla g_{\eta z}(U) \text{ and } R_z^{(v)} = \nabla g_{\eta z}(B_{\tau}(y_0)). \]
Let \( L \in \mathbb{N} \) be such that
\[ V_z^{(v)} \subseteq R_z^{(v)} \subseteq [-L, L]^r \quad (z \in \eta^{-1}B_{2\nu}(0), v \in B_{3\nu}(0)). \] (5.8)

We partition \( \mathbb{Z}^r \) into the following three subsets:
\[ H_z^{(v)}_1 = \{ h \in \mathbb{Z}^r : \frac{h}{j} \in V_z^{(v)} \}, \]
\[ H_z^{(v)}_2 = \{ h \in \mathbb{Z}^r : \text{dist}\left(\frac{h}{j}, V_z^{(v)}\right) \geq \rho \}, \]
and
\[ H_z^{(v)}_3 = \mathbb{Z}^r \setminus \left( H_z^{(v)}_1 \cup H_z^{(v)}_2 \right). \]

For each \( 1 \leq i \leq 3 \), we define
\[ M_i = \frac{1}{J} \sum_{1 \leq j \leq J} \left| \int_{\mathbb{R}^{d-r}} \int_{\mathbb{R}^{d-r}} \sum_{\frac{P}{2} \leq q \leq P} q^d \chi_{\eta}(z - \eta^{-1}v, v) \Phi(\eta q z) \sum_{h \in H_z^{(v)}} I_z^{(v)}(q; j; h) dz d\nu \right|. \]

Then (5.7) becomes
\[ \mathcal{M}(P, \alpha) \ll \frac{P^{d+1}}{J} + \eta^{-d} + M_1 + M_2 + M_3. \] (5.9)

We now bound each \( M_i \) separately.

5.1. **Case** \( h \in H_z^{(v)}_3 \). Let
\[ \xi(y) = \frac{j g_{\eta z}(y) - h \cdot y}{\text{dist}(h, jV_z^{(v)})} \text{ and } \beta = q \cdot \text{dist}(h, jV_z^{(v)}) \]
so that
\[ I_z^{(v)}(q; j; h) = \int_{\mathbb{R}^r} w(y)e(\beta \xi(y)) dy. \]

Then it follows from the definition of \( V_z^{(v)} \) that
\[ |\nabla \xi(y)| = \frac{|j \nabla g_{\eta z}(y) - h|}{\text{dist}(h, jV_z^{(v)})} \geq 1 \quad (y \in U). \]
Lemma 5.2. We have
\[ \|\xi\|_{C^\ell(U)} \ll 1, \]
where the implicit constant is independent of \( z, v, j \) and \( h \).

Proof. Let \( c > 0 \) be such that
\[
\sup_{y \in U, \, \, z \in B_{2\nu}(0), \, \, v \in B_{3\nu}(0)} |\nabla g_z(y)| < \frac{c}{2}.
\]
Suppose \( |h| \geq cj \). Then
\[
\left| \frac{h}{|h|} - \frac{j}{|h|} \right| |u| \geq 1 - \frac{j}{|h|} > \frac{1}{2} \quad (u \in V_z^{(v)}),
\]
hence
\[
\text{dist} \left( \frac{h}{|h|}, \frac{j}{|h|} V_z^{(v)} \right) \geq \frac{1}{2}.
\]
Therefore, it follows that
\[
|\xi(y)| = \left| \frac{j}{|h|} g_{jz}(y) - \frac{h}{|h|} y \right| \leq \frac{2|g_{jz}(y)|}{c} + 2|y| \ll 1 \quad (y \in U).
\]
On the other hand, suppose \( |h| < cj \). Then by the definition of \( H_z^{v,3} \) we have
\[
|\xi(y)| = \left| \frac{g_{jz}(y) - h}{j} \right| \leq \frac{|g_{jz}(y)| + c|y|}{\rho} \ll 1 \quad (y \in U).
\]
Thus we have proved
\[ \|\xi\|_{C^\ell(U)} \ll 1, \]
and the corresponding bounds for partial derivatives of \( \xi \) follow in a similar manner. \( \square \)

Recall \( 1 \leq r \leq 2 \) and \( \ell \geq 7 \). Therefore, by Lemma 3.4 it follows that
\[
\sum_{h \in \mathbb{H}_z^{(v)}} I_z^{(v)}(q; j; h) \ll q^{-\ell+1} \sum_{h \in \mathbb{H}_z^{(v)}} \text{dist}(h, jV_z^{(v)})^{-\ell+1} \quad (5.10)
\]
\[
\leq q^{-\ell+1} \sum_{m=0}^{\infty} \sum_{h \in \mathbb{Z}^{(v)}} \frac{1}{(2^m j \rho)^{\ell-1}} \sum_{2^m j \rho \leq \text{dist}(h, jV_z^{(v)}) < 2^{m+1} j \rho}
\]
\[
\ll q^{-\ell+1} \sum_{m=0}^{\infty} \frac{(jL + 2^{m+1} j \rho)^r}{(2^m j \rho)^{\ell-1}}
\]
\[
\ll q^{-\ell+1}.
\]
where the implicit constants are independent of \( z \) and \( v \). Consequently, by (4.13) and Lemma 5.1 we obtain

\[
M_3 \ll \frac{1}{J} \sum_{1 \leq j \leq J} \int_{\mathbb{R}^{d-r}} \sum_{\frac{P}{2} \leq q \leq P} q^{d-\ell+1} \int_{\mathbb{R}^{d-r}} \chi_{\eta}(z - \eta^{-1}v, v)|\Phi(\eta g z)|dz dv
\]

\[
= \int_{\mathbb{R}^{d-r}} \sum_{\frac{P}{2} \leq q \leq P} q^{r-\ell+1} \eta^{-d+r} \int_{\mathbb{R}^{d-r}} \chi_{\eta}\left(\eta^{-1}\left(\frac{z}{q} - v\right), v\right)|\Phi(z)|dz dv
\]

\[
\ll \int_{\mathbb{R}^{d-r}} \sum_{\frac{P}{2} \leq q \leq P} q^{r-\ell+1} \left(\frac{z}{q}\right)|\Phi(z)|dz
\]

\[
\ll \int_{\mathbb{R}^{d-r}} \sum_{\frac{P}{2} \leq q \leq P} q^{r-\ell+1} P^{d-r}(\log(P^2 + J))^{d-r}
\]

\[
\ll P^{d-r}(\log(P^2 + J))^{d-r}.
\]

5.2. Case \( h \in \mathbb{H}_{x;2}^{(v)} \). Let

\[
\varphi(y) = g_{y z}(y) - \frac{h}{j} \cdot y \quad \text{and} \quad \lambda = q_j
\]

so that

\[
I_{z}^{(v)}(q; j; h) = \int_{\mathbb{R}^{r}} w(y) e(\lambda \varphi(y))dy. \tag{5.12}
\]

It follows from the definition of \( \mathbb{H}_{x;2}^{(v)} \) and \( \rho \) that

\[
\frac{h}{j} \in R_{z}^{(v)}.
\]

We have

\[
\|\varphi\|_{C^\ell(U)} \leq \|g_{y z}\|_{C^\ell(U)} + \frac{|h|}{j}(2\nu + |y_0|) \leq \|g_{y z}\|_{C^\ell(U)} + L(2\nu + |y_0|) \lesssim 1,
\]

where the implicit constant is independent of \( z \) and \( v \). Since \( \nabla g_{y z} \) is a \( C^{\ell-1} \)-diffeomorphism on \( B_{\tau}(y_0) \), it follows that each \( h/j \in R_{z}^{(v)} \) determines

\[
y_{x;h/j} = y_{x;h/j}^{(v)} = (\nabla g_{y z})^{-1}\left(\frac{h}{j}\right) \in B_{\tau}(y_0), \tag{5.14}
\]
which is the unique critical point of \( \varphi \), i.e.

\[
0 = \nabla \varphi(y_{z,h/j}) = \nabla g_{\eta z}(y_{z,h/j}) - \frac{h}{j}.
\]  

(5.15)

Then by (3.1) we have

\[
\varphi(y_{z,h/j}) = g_{\eta z}(y_{z,h/j}) - \frac{h}{j} \cdot y_{z,h/j} = -g^*_{\eta z}\left(\frac{h}{j}\right),
\]  

(5.16)

which we will make use of in the next section.

**Lemma 5.3.** Suppose \( \tau \) is sufficiently small. Then we have

\[
\frac{|y - y_{z,h/j}|}{|\nabla \varphi(y)|} \ll 1 \quad (y \in U \setminus \{y_{z,h/j}\}),
\]

where the implicit constant is independent of \( z, v, j \) and \( h \).

**Proof.** Since

\[
\frac{|y - y_{z,h/j}|}{|\nabla \varphi(y)|} = \frac{|y - y_{z,h/j}|}{|\nabla g_{\eta z}(y) - \nabla g_{\eta z}(y_{z,h/j})|}
\]

and \( y_{z,h/j} \in B_\tau(y_0) \), it suffices to prove

\[
\frac{|y - t|}{|\nabla g_{\eta z}(y) - \nabla g_{\eta z}(t)|} \ll 1 \quad (y, t \in B_\tau(y_0), y \neq t),
\]  

(5.17)

where the implicit constant is independent of \( z, v, j \) and \( h \). By Taylor’s theorem, we obtain

\[
|\nabla g_{\eta z}(y) - \nabla g_{\eta z}(t) - H_{\eta z}(t) \cdot (y - t)| \ll |y - t|^2.
\]  

(5.18)

Recall that given an invertible \( d \times d \) real symmetric matrix \( A \), we have

\[
\lambda_{\text{min}}|t| \ll_d |A \cdot t| \quad (t \in \mathbb{R}^d),
\]

where \( \lambda_{\text{min}} \) is the minimum of the absolute values of the eigenvalues of \( A \). Therefore, from (1.6), (5.18) and the fact that the eigenvalues of a matrix are continuous in the entries of the matrix, it follows that there exist \( C, \kappa > 0 \) such that

\[
|y - t| \leq C|\nabla g_{\eta z}(y) - \nabla g_{\eta z}(t)|
\]

holds for all \( |y - t| \leq \kappa, \eta z \in B_{2\nu}(0) \) and \( v \in B_{3\nu}(0) \). Thus (5.17) holds provided \( \tau \) is sufficiently small with respect to \( \kappa \). \( \square \)

Since it follows from the definition of \( \mathbb{H}^{(v)}_{z;2} \) that

\[
\nabla g_{\eta z}(y_{z,h/j}) = \frac{h}{j} \not\in V_{z}^{(v)},
\]
we have \( y_{z,h,j} \notin U \), i.e. \( w(y_{z,h,j}) = 0 \). The cardinality of \( \mathbb{H}_{z,2}^{(v)} \) can be bounded as follows

\[
\# \mathbb{H}_{z,2}^{(v)} \leq \sum_{h \in \mathbb{Z}^r, \text{dist}(h/j, V_{z}^{(v)}) < \rho} 1 \leq \sum_{h \in \mathbb{Z}^r} 1 \ll j^r,
\]

where the implicit constant depends only on \( \rho \) and \( L \). Therefore, by Lemma 3.3 and (4.6) it follows that

\[
\sum_{h \in \mathbb{H}_{z,2}^{(v)}} I_z^{(v)}(q; j; h) \ll j^r \lambda^{-\frac{r}{q}} = q^{-\frac{r}{q}} j^{\frac{r}{q} - 1},
\]

where the implicit constant is independent of \( z \) and \( v \). Consequently, by (4.13) and Lemma 5.1 we obtain

\[
(5.19)
\]

\[
M_2 \ll \frac{1}{J} \sum_{1 \leq j \leq J} \int_{\mathbb{R}^{d-r}} \sum_{\frac{q}{j} \leq q \leq P} q^{d-\frac{r}{q} - 1} j^{\frac{r}{q} - 1} \int_{\mathbb{R}^{d-r}} \chi_\eta(z - \eta^{-1} v, v) |\Phi(\eta qz)| dz dv
\]

\[
\ll \int_{\mathbb{R}^{d-r}} \sum_{\frac{q}{j} \leq q \leq P} q^{\frac{r}{q} - 1} \int_{\mathbb{R}^{d-r}} \chi_\eta \left( \eta^{-1} \left( \frac{z}{q} - v \right), v \right) |\Phi(z)| dz dv
\]

\[
= \int_{\mathbb{R}^{d-r}} \sum_{\frac{q}{j} \leq q \leq P} q^{\frac{r}{q} - 1} \int_{\mathbb{R}^{d-r}} \chi \left( \frac{z}{q} \right) |\Phi(z)| dz
\]

\[
\ll P^{\frac{d-r}{q}} J^{-\frac{r}{q}} \int_{[-P,P]^{d-r}} |\Phi(z)| dz
\]

\[
\ll P^{d-r} \lambda^{1-r}(\log(P^2 + J))^{d-r}.
\]

**5.3 Case \( h \in \mathbb{H}_{z,1}^{(v)} \).** Let \( \varphi \) and \( \lambda \) be as in Section 5.2. In particular, we have (5.12), (5.13), (5.14), (5.15), (5.16) and Lemma 5.3. Since the eigenvalues of a matrix are continuous in the entries of the matrix, it follows from (4.6) that the signature of \( H_\varphi(y_{z,h,j}) = H_{\eta z}(y_{z,h,j}) \) is constant for all \( \eta z \in B_{3v}(0), \)

\( v \in B_{3\nu}(0), 1 \leq j \leq J \) and \( h \in \mathbb{H}_{z,1}^{(v)} \). Let us denote by \( \sigma \) the signature of \( H_\varphi(y_{z,h,j}) \). Then by Lemma 3.3 and (4.6) it follows that

\[
I_z^{(v)}(q; j; h) = \frac{w(y_{z,h,j})}{\sqrt{|\det H_{\eta z}(y_{z,h,j})|}} (qj)^{-\frac{r}{q}} e \left( qj g_{\eta z}(y_{z,h,j}) - qh \cdot y_{z,h,j} + \frac{\sigma}{8} \right)
\]

\[ + E_z^{(v)}(q; j; h), \]

where

\[
E_z^{(v)}(q; j; h) = O \left( ((qj)^{-\frac{r}{q}})^{\frac{1}{2}} \right)
\]

and the implicit constant is independent of \( z, v \) and \( h \).
Let
\[ W(u, z) = \frac{w((\nabla g_z)^{-1}(u))}{\sqrt{|\det H_{g_z}((\nabla g_z)^{-1}(u))|}}. \]

**Lemma 5.4.** Let \( z \in B_{2\nu}(0) \) and \( v \in B_{3\nu}(0) \). Then
\[ W(u, z) = 0 \quad \text{if} \quad u \notin \mathcal{X}, \]
where \( \mathcal{X} \) is defined in (4.9). We also have
\[ W \in C^1(\mathcal{X} \times B_{2\nu}(0)) \quad \text{and} \quad \sup_{u \in \mathcal{X}} \sum_{1 \leq i \leq d-r} \frac{\partial W}{\partial z_i}(u, z) \ll 1. \]

**Proof.** Since \( \text{supp} w \subseteq B_{2\nu}(y_0) \), the first statement is an immediate consequence of (4.6) and the definition of \( \mathcal{X} \). The second statement follows on recalling \( g \in C^{\ell-1}(B_\tau(y_0) \times B_{2\nu}(0) \times B_{3\nu}(0)) \), \( \nabla g_z \) is a \( C^{\ell-1} \)-diffeomorphism on \( B_\tau(y_0) \) for each \( z \in B_{2\nu}(0) \) and \( v \in B_{3\nu}(0) \), (4.10) and (4.11). \( \square \)

Therefore, with these notations and recalling (5.14) and (5.16) we obtain
\[ M_1 \ll \frac{1}{J} \sum_{1 \leq j \leq J} \sum_{h \in \mathcal{X}} M_0(P; j; h) + \mathcal{E}(P; J), \]
where
\[ M_0(P; j; h) = \int_{\mathbb{R}^{d-r}} \int_{\mathbb{R}^{d-r}} \left| \sum_{\frac{P}{2} \leq q \leq P} q^{d-\frac{r}{2}} \chi_{\eta}(z - \eta^{-1}v, v) \Phi(\eta q z) W \left( \frac{h}{j}, \eta z \right) e \left( -q j g^*_\eta \left( \frac{h}{j} \right) \right) \right| dz dv \]
and
\[ \mathcal{E}(P; J) = \frac{1}{J} \sum_{1 \leq j \leq J} \int_{\mathbb{R}^{d-r}} \int_{\mathbb{R}^{d-r}} \sum_{\frac{P}{2} \leq q \leq P} q^d \chi_{\eta}(z - \eta^{-1}v, v) \]
\[ \sum_{h \in \mathbb{H}_{z,1}^{(v)}} |\Phi(\eta q z) E_z^{(v)}(q; j; h)| dz dv. \]

The cardinality of \( \mathbb{H}_{z,1}^{(v)} \) can be bounded as follows
\[ \# \mathbb{H}_{z,1}^{(v)} \leq \sum_{h \in \mathbb{Z}^r \atop \frac{h}{j} \in [-L,L]^r} 1 \ll j^r, \]
where the implicit constant depends only on \( L \). Thus it follows as in (5.19) that

\[
\mathcal{E}(P; J) \leq \frac{1}{J} \sum_{1 \leq j \leq J} \int_{\mathbb{R}^{d-r}} \left| q^d J^r(q_j)^{\frac{1}{2} - 1} \int_{\mathbb{R}^{d-r}} \chi(q_j z - \eta^{-1} v, v) |\Phi(\eta q z)| d z d v \right|
\]

\[
\leq J^{\frac{d-r}{2}} \int_{\mathbb{R}^{d-r}} \sum_{\frac{1}{2} \leq q \leq P} q^{\frac{d-r}{2} - 1} \eta^{-d+r} \int_{\mathbb{R}^{d-r}} \chi(q^{-1} \left( \frac{z}{q} - v \right), v) |\Phi(z)| d z d v
\]

\[
\leq P^{d-r} J^{\frac{d-r}{2}} (\log(P^2 + J))^{d-r}.
\]

It is clear from the definition of \( \chi \) that

\[
\chi(q^{-1} \left( \frac{z}{q} - \eta^{-1} v \right), v) \neq 0
\]

implies

\[
z \in \eta^{-1} q \text{ supp } \chi \subseteq \eta^{-1} P B_{2q}(0) \quad \text{and} \quad v \in B_{3v}(0).
\]

Since it follows from (4.4) that

\[
q_j g^*_{q_j/q}(u) = q_j G^*(u) + \eta j \sum_{1 \leq i \leq d-r} \frac{\partial \mathcal{F}}{\partial z_i}(u, v) z_i
\]

holds for all \( u \in B_{\mu}(u_0), \eta z/q \in B_{2q}(0) \) and \( v \in B_{3v}(0) \), we have

\[
M_0(P; j; h) = \int_{B_{3u}(0)} \int_{[-\eta^{-1} P, \eta^{-1} P]^{d-r}} \left| \sum_{\frac{1}{2} \leq q \leq P} q^2 \chi(q_j z - \eta^{-1} v, v) \Phi(\eta z) W \left( \frac{h}{j}, \frac{\eta z}{q} \right) e \left( -q_j G^*(\eta z) \left( \frac{h}{j} \right) - \eta j \sum_{1 \leq i \leq d-r} \frac{\partial \mathcal{F}}{\partial z_i}(u, v) z_i \right) \right| d z d v
\]

\[
\leq \int_{[-\eta^{-1} P, \eta^{-1} P]^{d-r}} |\Phi(\eta z)| |\Gamma_z(P; j; h)| d z,
\]

where

\[
\Gamma_z(P; j; h) = \int_{B_{3u}(0)} \left| \sum_{\frac{1}{2} \leq q \leq P} q^2 \chi(q_j z - \eta^{-1} v, v) W \left( \frac{h}{j}, \frac{\eta z}{q} \right) e \left( -q_j G^*(\eta z) \left( \frac{h}{j} \right) \right) \right| d v.
\]

Let

\[
\varphi(z) = \chi(z) W \left( \frac{h}{j}, z \right) \quad \text{and} \quad \beta(v) = -j G^*(\frac{h}{j}).
\]
By partial summation and Lemma 5.4 we have
\[ \sum_{\frac{q}{2} \leq q \leq P} q^2 \chi_{\eta} \left( \frac{z}{q} - \eta^{-1} v, v \right) W \left( \frac{h_j}{j}, \frac{\eta z}{q} \right) \left( -q j \mathcal{G}_0 \left( \frac{h_j}{j} \right) \right) \]
\[ \leq \sum_{\frac{q}{2} \leq q \leq P} q^2 \prod_{1 \leq i \leq d-r} \omega_0 \left( \frac{z_i}{q} - \eta^{-1} v_i \right) \cdot \varphi \left( \frac{\eta z}{q} \right) e(q \beta(v)) \]
\[ \ll P^2 \min(P, \|\beta(v)\|^{-1}) \left( \prod_{1 \leq i \leq d-r} \omega_0 \left( \frac{z_i}{P} - \eta^{-1} v_i \right) + \prod_{1 \leq i \leq d-r} \omega_0 \left( \frac{2z_i}{P} - \eta^{-1} v_i \right) \right) \]
\[ + \min(P, \|\beta(v)\|^{-1}) \int_{P/2}^P u^{\frac{r}{2}} \prod_{1 \leq i \leq d-r} \omega_0 \left( \frac{z_i}{u} - \eta^{-1} v_i \right) \, du \]
\[ + \min(P, \|\beta(v)\|^{-1}) \int_{P/2}^P u^{\frac{r}{2}} \sum_{1 \leq i \leq d-r} \left| \frac{z_i}{u^2} \omega_0' \left( \frac{z_i}{u} - \eta^{-1} v_i \right) \right| \prod_{1 \leq j \leq d-r} \omega_0 \left( \frac{z_j}{u} - \eta^{-1} v_j \right) \, du \]
\[ + \min(P, \|\beta(v)\|^{-1}) \int_{P/2}^P u^{\frac{r}{2}} \prod_{1 \leq i \leq d-r} \omega_0 \left( \frac{z_i}{u} - \eta^{-1} v_i \right) \cdot \frac{\eta|z|}{u^2} 1_{B_{2\nu}(0)} \left( \frac{\eta z}{u} \right) \, du. \]

Let \( P/2 \leq u \leq P \). Then
\[ \int_{\mathbb{R}^{d-r}} \prod_{1 \leq i \leq d-r} \omega_0 \left( \frac{z_i}{u} - \eta^{-1} v_i \right) \, dv \]
In the situation where $\mathcal{S}_+$ or $\mathcal{S}_-$ is an empty set, we set $a_1 = a_2 = P$ or $b_1 = b_2 = P$ respectively. With these notations, it follows that

$$
\int_{R} \int_{P/2}^{P} \left| \frac{z_i}{u^2} \omega'_0 \left( \frac{z_i}{u} - v_i \right) \right| \, du \, dv_i
= \int_{R} \left( \int_{\mathcal{S}_+} \text{sign}(z_i) \frac{z_i}{u^2} \omega'_0 \left( \frac{z_i}{u} - v_i \right) \, du + \int_{\mathcal{S}_-} -\text{sign}(z_i) \frac{z_i}{u^2} \omega'_0 \left( \frac{z_i}{u} - v_i \right) \, du \right) \, dv_i
= \int_{R} \left| \omega_0 \left( \frac{z_i}{a_2} - v_i \right) - \omega_0 \left( \frac{z_i}{a_1} - v_i \right) \right| + \left| \omega_0 \left( \frac{z_i}{b_2} - v_i \right) - \omega_0 \left( \frac{z_i}{b_1} - v_i \right) \right| \, dv_i
\leq \int_{R} \omega_0 \left( \frac{z_i}{a_2} - v_i \right) + \omega_0 \left( \frac{z_i}{a_1} - v_i \right) + \omega_0 \left( \frac{z_i}{b_2} - v_i \right) + \omega_0 \left( \frac{z_i}{b_1} - v_i \right) \, dv_i
= 4.
$$

Then from this inequality we obtain

$$
\int_{R_{d-r}} \int_{P/2}^{P} u^2 \left| \frac{z_i}{u^2} \omega'_0 \left( \frac{z_i}{u} - \eta^{-1} v_i \right) \right| \prod_{1 \leq j \leq d-r \atop j \neq i} \omega_0 \left( \frac{z_j}{u} - \eta^{-1} v_j \right) \, du \, dv
= \eta^{d-r} \int_{R_{d-r}} \int_{P/2}^{P} u^2 \left| \frac{z_i}{u^2} \omega'_0 \left( \frac{z_i}{u} - v_i \right) \right| \prod_{1 \leq j \leq d-r \atop j \neq i} \omega_0 \left( \frac{z_j}{u} - v_j \right) \, du \, dv_i
= \eta^{d-r} \int_{R} \int_{P/2}^{P} u^2 \left| \frac{z_i}{u^2} \omega'_0 \left( \frac{z_i}{u} - v_i \right) \right| \, du \, dv_i
\leq \eta^{d-r} P^2 \int_{R} \int_{P/2}^{P} \left| \frac{z_i}{u^2} \omega'_0 \left( \frac{z_i}{u} - v_i \right) \right| \, du \, dv_i
\ll \eta^{d-r} P^2.
$$
Therefore, by combining (5.23), (5.24) and (5.25) we obtain
\[
\Gamma_\zeta(P; j; h) = \int_{B_{3\nu}(0)} \left| \sum_{\eta < q \leq P} q^{\frac{r}{2}} \chi_\eta \left( \frac{z}{q} - \eta^{-1} v, v \right) W \left( \frac{h}{j}, \eta z \right) e(q\beta(v)) \right| dv
\]
\[
\ll \eta^{d-r} P^{\frac{r}{2}} \sup_{v \in B_{3\nu}(0)} \min(P, \|\beta(v)\|^{-1})
\]
\[
+ \eta^{d-r} \int_{P/2}^{P} \left( u^{\frac{r}{2}} - 1 + u^{\frac{r}{2}} \eta |z| \right) \, \frac{1}{u^2} \, dz \sup_{v \in B_{3\nu}(0)} \min(P, \|\beta(v)\|^{-1})
\]
\[
\ll \eta^{d-r} P^{\frac{r}{2}} \sup_{v \in B_{3\nu}(0)} \min(P, \|\beta(v)\|^{-1}),
\]
where the implicit constant is independent of \(z, j\) and \(h\). Then it follows by (5.22), Lemma 5.1 and this estimate that
\[
M_0(P; j; h) \ll \int_{[-\eta^{-1} p, \eta^{-1} p]^{d-r}} |\Phi(\eta z)| \Gamma_\zeta(P; j; h) \, dz
\]
\[
\ll \eta^{d-r} P^{\frac{r}{2}} \sup_{v \in B_{3\nu}(0)} \min(P, \|jG^\star(h/j)\|^{-1}) \int_{[-\eta^{-1} p, \eta^{-1} p]^{d-r}} |\Phi(\eta z)| \, dz
\]
\[
= \int_{[-\eta^{-1} p, \eta^{-1} p]^{d-r}} |\Phi(z)| \, dz
\]
\[
\ll P^{d-r} (\log(P^2 + J))^{d-r} \sup_{v \in B_{3\nu}(0)} \min(P, \|jG^\star(h/j)\|^{-1}).
\]

Therefore, by substituting (5.21) and this estimate into (5.20) we obtain
\[
M_1 \ll \frac{P^{d-r}}{J} (\log(P^2 + J))^{d-r} \sup_{v \in B_{3\nu}(0)} \sum_{1 \leq j \leq J} j^{-\frac{r}{2}} \sum_{h \in \mathcal{X}} \min(P, \|jG^\star(h/j)\|^{-1})
\]
\[
+ P^{d-r} J^{\frac{r}{2}-1} (\log(P^2 + J))^{d-r}.
\]

5.4. Final estimate. We collect together the estimates obtained in this section, namely (5.11), (5.19) and (5.27), and substitute them into (5.9) to obtain
\[
\mathcal{M}(P, \alpha)
\]
\[
\ll \frac{P^{d+1}}{J} + \eta^{-d} + P^{d-r} (\log(P^2 + J))^{d-r} + P^{d-r} J^{\frac{r}{2}-1} (\log(P^2 + J))^{d-r}
\]
\[
+ \frac{P^{d-r}}{J} (\log(P^2 + J))^{d-r} \sup_{v \in B_{3\nu}(0)} \sum_{0 \leq s \leq \frac{\log J}{\log \alpha}} \left( \frac{2^s}{J} \right)^{\frac{r}{2}} \sum_{1 \leq j \leq J} \sum_{h \in \mathcal{X}} \min(P, \|jG^\star(h/j)\|^{-1}).
\]
Let $\gamma > 0$ be sufficiently small and $W \in C^\infty_c(\mathbb{R}^r)$ such that
\[ 1_{\mathcal{X}}(u) \leq W(u) \quad (u \in \mathbb{R}^r) \quad \text{and} \quad \text{dist}(\partial(\text{supp } W), \partial \mathcal{X}) \leq \gamma. \]

In particular, by combining these with Lemmas 4.2 and 4.3 we see that the hypotheses of Proposition 3.6 (with $G^\ast v$) are satisfied. Therefore, by Proposition 3.6 we obtain
\[
\sup_{v \in B_{3\nu}(0)} \sum_{0 \leq s \leq \log J} \left( \frac{J}{2^s} \right)^{\frac{r}{2}} \sum_{J \leq j \leq 2^s} \min(P, \| j G^\ast_v(h/j) \|^{-1})
\leq \sup_{v \in B_{3\nu}(0)} \sum_{0 \leq s \leq \log J} \left( \frac{J}{2^s} \right)^{\frac{r}{2}} \sum_{J \leq j \leq 2^s} W(h/j) \min(P, \| j G^\ast_v(h/j) \|^{-1})
\leq P^e \sum_{0 \leq s \leq \log J} \left( \frac{J}{2^s} \right)^{\frac{r}{2}} + PJ^e \sum_{0 \leq s \leq \log J} \left( \frac{J}{2^s} \right)^{\frac{r}{2}}
\ll (J^{-\frac{r}{2}} + PJ^e) P^e J^e.
\]

Finally, we substitute this bound into (5.28) and arrive at
\[
\mathcal{M}(P, \alpha) \ll \frac{P^{d+1}}{J} + \eta^{-d} + P^{d-r} (\log(P^2 + J))^{d-r} + P^{d-\frac{r}{2}} J^{\frac{r}{2}-1} (\log(P^2 + J))^{d-r}
+ (P^{d-\frac{r}{2}} J^{\frac{r}{2}} + P^{d+1-\frac{r}{2}} J^{-\frac{r}{2}-1}) P^e J^e
\ll \alpha P^{d+1} + \eta^{-d} + (P^{d-\frac{r}{2}} \alpha^{\frac{r}{2}} + P^{d+1-\frac{r}{2}} \alpha^{-\frac{r}{2}}) P^e \alpha^{-e},
\]
which completes the proof of Proposition 4.4.

**Appendix A. Differential Geometry by Matthew Beckett**

In this appendix, by manifolds we mean connected real $C^\ell$-manifolds for any $\ell \in \mathbb{Z}_{\geq 2} \cup \{\infty\}$. For the majority of the appendix, we will consider Riemannian submanifolds $X \subseteq \mathbb{R}^n$, which is to say that we equip an embedded submanifold $X$ with the metric induced from the Euclidean metric on $\mathbb{R}^n$. This is the unique metric on $X$ for which the inclusion $\iota : X \hookrightarrow \mathbb{R}^n$ is an isometry. Given such a submanifold $X \subseteq \mathbb{R}^n$, by the implicit function theorem without loss of generality we can give a local description of $X$ as
\[
X = \{(x_1, \ldots, x_d, f_1, \ldots, f_{n-d}) : x_1, \ldots, x_d \in \mathbb{R}\}, \quad \text{(A.1)}
\]
where $f_1, \ldots, f_{n-d} \in C^\ell(\mathbb{R}^d)$ depend on the variables $x_1, \ldots, x_d$. In this case, $X$ is $d$-dimensional, and the variables $x_1, \ldots, x_d$ give local coordinates for $X$. 
Definition A.1. Given a local description such as in (A.1) around a point \( p \), we say that \( p \) is a bad point if
\[
\frac{\partial^2 f_i}{\partial x_j \partial x_k} = 0
\]
for all \( 1 \leq i \leq n - d \) and \( 1 \leq j, k \leq d \).

In Proposition A.2, it will become apparent that \( p \) being a bad point is independent of the choice of a local representation.

A.1. Second Fundamental Form. At a point \( p \) of a Riemannian submanifold \( X \subseteq \mathbb{R}^n \), the tangent space of the ambient Euclidean space decomposes orthogonally as \( T_p \mathbb{R}^n = T_p X \oplus N_p \), where we use \( N_p \) to denote the orthogonal complement of the tangent space \( T_p X \).

Let \( V \) and \( W \) be vector fields on \( X \), extended to some neighbourhood of \( p \) in \( \mathbb{R}^n \). Let \( D_V W \) denote the derivative of \( W \) in the direction \( V \). Taking the orthogonal decomposition of \( T_p \mathbb{R}^n \), we write
\[
(D_V W)_p = \nabla_{V_p} W + II_p(V, W),
\]
where \( \nabla_{V_p} W \) denotes the orthogonal projection of \( (D_V W)_p \) onto \( T_p X \) and \( II_p(V, W) \) denotes the projection onto \( N_p \). It is a classical result that \( \nabla \) is then the Levi-Civita connection on \( X \), while \( II \) is symmetric in \( V \). We recall that \( II \) is called the second fundamental form at \( p \), denoted by \( II \) when we allow \( p \) to vary. Although \( II \) was defined in terms of some extensions of \( V \) and \( W \), it can be shown that \( II \) is in fact independent of the extensions chosen, and thus is well-defined on tangent vectors of \( X \).

To relate the second fundamental form with a description of bad points, we work with the local description of \( X \) as in (A.1). For \( 1 \leq i \leq d \), let \( V_i \) denote the vector field on \( \mathbb{R}^n \) defined by
\[
V_i = \left( 0, \ldots, 0, 1, 0, \ldots 0, \frac{\partial f_1}{\partial x_i}, \ldots, \frac{\partial f_n - d}{\partial x_i} \right),
\]
where the 1 occurs in the \( i \)-th component. Then the set of vectors \( \{(V_1)_q, \ldots, (V_d)_q\} \) is a basis for \( T_q X \) for each \( q \in X \), and thus the vector fields \( V_i \) define a local tangent frame for \( TX \) when restricted to \( X \).

For \( 1 \leq k \leq n - d \), we also define a vector field \( W_k \) by
\[
W_k = \left( -\frac{\partial f_k}{\partial x_1}, \ldots, -\frac{\partial f_k}{\partial x_d}, 0, \ldots, 0, 1, 0, \ldots, 0 \right),
\]
where the 1 occurs in the \( (d + k) \)-th component. Each \( W_k \) is orthogonal to each \( V_i \), and so \( (W_k)_q \in N_q \). In fact, by dimension count we see that \( \{(W_1)_q, \ldots, (W_{n-d})_q\} \) is a basis for \( N_q \). Therefore, we now have bases for the
two orthogonal components in the decomposition $T_p\mathbb{R}^n = T_pX \oplus N_p$, although we note that neither basis is an orthogonal basis of its respective space.

We now compute the derivatives $D_{V_i}V_j$. The derivative in the direction $V_i$ is simply

$$\frac{\partial}{\partial x_i} + \sum_{1 \leq h \leq n-d} \frac{\partial f_h}{\partial x_i} \frac{\partial}{\partial x_{d+h}}.$$  

Since the components of $V_j$ are constant in $x_{d+1}, \ldots, x_n$, we then see that

$$D_{V_i}V_j = \left(0, \ldots, 0, \frac{\partial^2 f_1}{\partial x_i \partial x_j}, \ldots, \frac{\partial^2 f_{n-d}}{\partial x_i \partial x_j}\right).$$  

From this expression it is clear that $p$ is a bad point if and only if $(D_{V_i}V_j)_p = 0$ for all $1 \leq i, j \leq d$. In particular, if $p$ is a bad point, then $\Pi_p = 0$, where we note that as a tensor $\Pi_p = 0$ means $\Pi_p(V, W) = 0$ for all vectors $V, W \in T_pX$.

Conversely, suppose $p$ is not a bad point. Then $(D_{V_i}V_j)_p$ is non-zero for some $1 \leq i, j \leq d$, and because its first $d$ components are 0, it cannot be expressed as a linear combination solely of the $V_i$. Therefore, it must have some non-zero component in $N_p$, and hence $\Pi_p(V_i, V_j) \neq 0$. Since this is true for arbitrary basis vectors $V_i, V_j$, we have established the following proposition.

**Proposition A.2.** With $X \subseteq \mathbb{R}^n$ as above, the point $p \in X$ is a bad point if and only if $\Pi_p = 0$.

Of particular note, the second fundamental form depends on the embedding used rather than the particular choice of coordinates. As such, to apply Proposition A.2 one may choose any coordinates for which $\Pi_p$ is simple to compute, not merely coordinates of the form (A.1).

**A.2. Curvature.** By relating the second fundamental form with various notions of curvature, we can obtain conditions that guarantee that points will not be bad. For a moment consider the more general case of an isometric embedding of Riemannian manifolds $X \subseteq M$, with respective Riemann curvature tensors denoted $R_X$ and $R_M$. As in the previous section, let $\Pi$ denote the second fundamental form. Then writing the Riemannian metric on $M$ as $g_M(\cdot, \cdot) = \langle \cdot, \cdot \rangle$, the Gauss equation states

$$\langle R_M(V_1, V_2)V_3, V_4 \rangle = \langle R_X(V_1, V_2)V_3, V_4 \rangle + \langle \Pi(V_1, V_3), \Pi(V_2, V_4) \rangle \tag{A.2}$$

$$- \langle \Pi(V_2, V_3), \Pi(V_1, V_4) \rangle.$$  

In our case of interest, we let $M = \mathbb{R}^n$ and so $R_M = 0$. For ease of notation, we will write $R = R_X$. Then (A.2) becomes

$$0 = \langle R(V_1, V_2)V_3, V_4 \rangle + \langle \Pi(V_1, V_3), \Pi(V_2, V_4) \rangle \tag{A.3}$$

$$- \langle \Pi(V_2, V_3), \Pi(V_1, V_4) \rangle.$$  

From this we obtain various corollaries to Proposition A.2.
Corollary A.3. If the tensor $R_p \neq 0$, then $p$ is not a bad point.

Proof. If $R_p$ is non-zero, then the first term of (A.3) is non-zero for some choice of $V_1, V_2, V_3, V_4$. Therefore, there must be at least one other non-zero term, and hence $\Pi_p \neq 0$. By Proposition A.2, $p$ is not a bad point. \[ \square \]

We then obtain the following corollaries by taking successive traces of $R_p$ to evaluate the Ricci curvature $\text{Ric}$ and the scalar curvature $S$.

Corollary A.4. If $\text{Ric}_p \neq 0$, then $p$ is not a bad point.

Proof. Since $\text{Ric}_p$ is obtained as a trace of $R_p$, if $\text{Ric}_p \neq 0$ then $R_p \neq 0$. \[ \square \]

Corollary A.5. If $S_p \neq 0$, then $p$ is not a bad point.

Proof. Since $S_p$ is obtained as a trace of $\text{Ric}_p$, if $S_p \neq 0$ then $\text{Ric}_p \neq 0$. \[ \square \]

The above notions of curvature are all intrinsic notions to the manifold $X$, although it should be noted that using the Gauss equation requires using the metric induced by the embedding $X \subseteq \mathbb{R}^n$. For an extrinsic notion of curvature, we recall the definition of mean curvature

$$H_p = -\frac{1}{\dim X} \text{tr} \Pi_p,$$

(A.4)
a trace of the second fundamental form, which is a vector normal to the manifold at $p$. Since a tensor that vanishes must have a trace that vanishes, we then obtain one more corollary to Proposition A.2.

Corollary A.6. If $H_p \neq 0$, then $p$ is not a bad point.

References

[1] V. Beresnevich, *Rational points near manifolds and metric Diophantine approximation*, Ann. of Math. (2) 175 (2012), no. 1, 187–235.

[2] V. Beresnevich, D. Dickinson and S. Velani, *Diophantine approximation on planar curves and the distribution of rational points*, with Appendix II by R.C. Vaughan. Ann. of Math. (2) 166 (2007), no. 2, 367–426.

[3] V. Beresnevich, R.C. Vaughan, S. Velani and E. Zorin, *Diophantine approximation on manifolds and the distribution of rational points: Contribution to the convergence theory*, Int. Math. Res. Not. IMRN 2017, no.10, 2885–2908.

[4] T.D. Browning, D.R. Heath-Brown and P. Salberger, *Counting rational points on algebraic varieties*, Duke Math. J. 132 (2006), 545–578.

[5] W. Castryck, R. Cluckers, P. Dittmann and K.H. Nguyen, *The dimension growth conjecture, polynomial in the degree and without logarithmic factors*, Algebra Number Theory 14 (2020), no. 8, 2261–2294.

[6] D.R. Heath-Brown, *A new form of the circle method, and its application to quadratic forms*. J. reine angew. Math. 481 (1996), 149–206.

[7] D.R. Heath-Brown, *The density of rational points on curves and surfaces*, Ann. of Math. (2) 155 (2002), no. 2, 553–595.
[8] L. Hörmander, *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis*, Second edition. Grundlehren der Mathematischen Wissenschaften, 256. Springer-Verlag, Berlin, 1990.

[9] J.-J. Huang, *The density of rational points near hypersurfaces*, Duke Math. J. 169 (2020), 2045–2077.

[10] M.N. Huxley, *The rational points close to a curve*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 21 (1994), no. 3, 357–375.

[11] P. Salberger, *Counting rational points on projective varieties*, (2013), preprint.

[12] J.-P. Serre, *Topics in Galois theory*, Research Notes in Mathematics, vol. 1, Jones and Bartlett Publishers, Boston, MA, 1992, Lecture notes prepared by Henri Damon, with a foreword by Darmon and the author.

[13] R.C. Vaughan and S. Velani, *Diophantine approximation on planar curves: the convergence theory*, Invent. Math. 166 (2006), no. 1, 103–124.

Mathematisch Instituut, Universiteit Utrecht, Budapestlaan 6, NL-3584CD Utrecht, The Netherlands

Email address: s.yamagishi@uu.nl