DETERMINANT VALUES ON LATTICES

WOOYEON KIM AND HEE OH

Abstract. We study the distribution of determinant values on lattices of $\mathrm{M}_d(\mathbb{R})$. Let $d \geq 2$. Let $\Lambda < \mathrm{M}_d(\mathbb{R})$ be a lattice whose elements all have algebraic entries. We prove that, if the determinant values on Λ are not contained in a scalar multiple of $\mathbb{Z}$, then for every $a < b$,

$$\#\{v \in \Lambda : \|v\| < T,\ a < \det v < b,\ \det v \neq 0\} \sim \frac{C_d}{\mathrm{covol}(\Lambda)}(b-a)T^{d(d-1)}$$

as $T \to \infty$, where $\|\cdot\|$ denotes the Frobenius norm and $C_d > 0$ is a constant depending only on d. Under a mild additional hypothesis, which is automatic for $d = 2, 3$, we also prove an asymptotic for determinant-zero lattice points. All of these results hold more generally for Diophantine lattices.

The case $d = 2$ recovers the Eskin–Margulis–Mozes theorem in the quantitative Oppenheim problem for quadratic forms of signature $(2, 2)$.

1. Introduction

The distribution of values of a polynomial $F(x_1, \ldots, x_N)$ on lattice points is a classical problem in number theory. For indefinite quadratic forms, this circle of questions has been extensively studied under the name of the quantitative Oppenheim problem; see [26, 27, 10, 13, 14, 21]. Although higher-degree forms in many variables are often accessible by the Hardy–Littlewood circle method, the determinant lies far outside the usual Birch range: for a degree-d form, Birch-type hypotheses require roughly $d2^d$ nonsingular directions, whereas the determinant has only d^2 variables and its singular locus has codimension 4.

In this paper we study the distribution of determinant values on lattices in $\mathrm{M}_d(\mathbb{R})$. A lattice in $\mathrm{M}_d(\mathbb{R})$ means a discrete subgroup of full rank, equivalently the $\mathbb{Z}$-span of d^2 linearly independent matrices. Given such a lattice Λ, we study the distribution in $\mathbb{R}$ of the determinant values

$$\det(\Lambda) := \{\det v : \ v \in \Lambda\}.$$

Fix a maximal compact subgroup $K_0 < \mathrm{SL}_d(\mathbb{R})$ and a norm $\|\cdot\|$ on $\mathrm{M}_d(\mathbb{R})$ that is invariant under left and right multiplication by K_0; the Frobenius norm is the basic example. For a lattice $\Lambda < \mathrm{M}_d(\mathbb{R})$, real numbers $a < b$, and $T > 0$, set

$$N_\Lambda(a, b; T) := \#\{v \in \Lambda : \ \|v\| < T,\ a < \det v < b\},$$

Oh is partially supported by the NSF grant No. DMS-2450703.

and

$$N^{\times}_{\Lambda}(a,b;T) := \#\{v \in \Lambda : \ \|v\| < T, \ a < \det v < b, \ \det v \neq 0\}.$$

The distinction is essential when $0 \in (a,b)$: determinant-zero lattice points may contribute on the same scale as the nonsingular main term.

1.1. **The nonsingular counting theorem.** Our first main result gives a precise asymptotic for $N^{\times}_{\Lambda}(a,b;T)$.

Theorem 1.1. *Let $\Lambda < \mathrm{M}_d(\mathbb{R})$ be a lattice whose elements have algebraic entries, and assume that*

$$\det(\Lambda) \not\subset \lambda\mathbb{Z} \qquad \textit{for every } \lambda \in \mathbb{R}^{\times}. \tag{1.1}$$

Then there exists a constant $C_{\|\cdot\|} > 0$, depending only on the norm $\|\cdot\|$, such that, for every $a < b$,

$$N^{\times}_{\Lambda}(a,b;T) \sim \frac{C_{\|\cdot\|}}{\mathrm{covol}(\Lambda)}(b-a)T^{d(d-1)} \qquad (T \to \infty). \tag{1.2}$$

Denote by vol the Lebesgue measure on $\mathrm{M}_d(\mathbb{R})$. All covolumes are computed with respect to this same measure. The constant $C_{\|\cdot\|}$ is characterized by the volume asymptotic

$$\mathrm{vol}\{v \in \mathrm{M}_d(\mathbb{R}) : \ \|v\| < T, \ a < \det v < b\} \sim C_{\|\cdot\|}(b-a)T^{d(d-1)}, \tag{1.3}$$

proven in Lemma 12.16. Therefore Theorem 1.1 may also be written as

$$N^{\times}_{\Lambda}(a,b;T) \sim \frac{1}{\mathrm{covol}(\Lambda)}\,\mathrm{vol}\{v \in \mathrm{M}_d(\mathbb{R}) : \ \|v\| < T, \ a < \det v < b\}.$$

1.2. **The singular contribution.** The exclusion of the zero determinant in Theorem 1.1 is necessary. Indeed, if all columns of a matrix are constrained to lie in a fixed proper subspace $U < \mathbb{R}^d$, then the determinant vanishes identically; the same is true if all rows are constrained to lie in such a subspace. We call such subspaces column- and row-isotropic, respectively. When $\dim U = d-1$, they have dimension $d(d-1)$, exactly the exponent appearing in Theorem 1.1. Thus a $\mathbb{Z}$-submodule of Λ of rank $d(d-1)$ contained in an isotropic subspace may contribute on the same scale as the regular main term.

The isotropic noncoincidence condition used below rules out certain quotient-lattice commensurabilities between distinct maximal Λ-rational column-isotropic subspaces, and separately between distinct maximal Λ-rational row-isotropic subspaces, that could create additional main-order overlaps in the singular contribution. The precise definition appears in Definition 3.14; in particular, the condition is vacuous for $d = 2, 3$. For examples, see Example 13.6.

Theorem 1.2. *Let Λ be as in Theorem 1.1. Suppose either that $d \leq 3$, or that Λ satisfies the isotropic noncoincidence condition. Then the limit*

$$c^{\mathrm{sing}}_{\Lambda} := \lim_{T\to\infty} T^{-d(d-1)}\#\{v \in \Lambda : \ \|v\| < T, \ \det v = 0\} \tag{1.4}$$

exists and is finite. Moreover, $c_\Lambda^{\mathrm{sing}} > 0$ if and only if Λ contains a rank-$d(d-1)$ submodule on which the determinant vanishes identically.

Consequently, for every $a < b$,

$$N_\Lambda(a, b; T) \sim \left(\frac{C_{\|\cdot\|}}{\operatorname{covol}(\Lambda)} (b - a) + c_\Lambda^{\mathrm{sing}} \mathbf{1}_{\{0 \in (a,b)\}} \right) T^{d(d-1)}. \tag{1.5}$$

The irrationality hypothesis (1.1) in Theorem 1.1 is easy to test: it is equivalent to the existence of $v, w \in \Lambda$, with $\det w \neq 0$, such that

$$\frac{\det v}{\det w} \notin \mathbb{Q}.$$

For example, when $d = 3$, both theorems apply to

$$\Lambda_{\sqrt{2}} := \begin{pmatrix} \sqrt{2}\,\mathbb{Z} & \mathbb{Z} & \mathbb{Z} \\ \mathbb{Z} & \mathbb{Z} & \mathbb{Z} \\ \mathbb{Z} & \mathbb{Z} & \mathbb{Z} \end{pmatrix}, \tag{1.6}$$

since $\det(\Lambda_{\sqrt{2}})$ contains both a nonzero rational number and a nonzero rational multiple of $\sqrt{2}$. The matrices with first row zero form a rank-six singular submodule, so $c_{\Lambda_{\sqrt{2}}}^{\mathrm{sing}} > 0$.

Moreover, $\Lambda_{\sqrt{2}}$ contains infinitely many rank-six submodules on which the determinant vanishes identically. Writing v_1, v_2, v_3 for the columns of $v \in \Lambda_{\sqrt{2}}$, each primitive pair $(p, q) \in \mathbb{Z}^2$, modulo sign, defines such a submodule:

$$\{v \in \Lambda_{\sqrt{2}} : pv_2 + qv_3 = 0\}.$$

Thus the singular contribution in Theorem 1.2 is not merely a finite-exceptional-set phenomenon: an irrational algebraic lattice may contain infinitely many rank-$d(d-1)$ singular submodules, yet their aggregate contribution still has a finite $T^{d(d-1)}$-normalized limit. By contrast, the lattice constructed from a cubic division algebra in Example 13.5 contains no such submodule and satisfies $c_\Lambda^{\mathrm{sing}} = 0$.

1.3. **Diophantine lattices.** The algebraicity assumption on Λ in the above theorems can be replaced by a Diophantine condition on rational subspaces, formulated in terms of their Plücker vectors; see Definition 3.6 and Theorem 3.8. Informally, this condition requires rational subspaces to remain polynomially separated from the exceptional column- and row-isotropic directions unless they lie in those directions exactly. Some arithmetic separation is necessary. Katznelson proved that the number of rank-$(d-1)$ matrices in $\mathrm{M}_d(\mathbb{Z})$ of norm at most T is of order

$$T^{d(d-1)} \log T;$$

see [20]. The logarithmic factor reflects the failure of summability on the critical $T^{d(d-1)}$-scale in the split rational lattice. In the present setting, the Diophantine condition prevents non-isotropic rational subspaces from having Plücker vectors too close to the exceptional isotropic directions.

Exact isotropic subspaces, which may occur in infinite families, are treated separately in Theorem 3.15.

1.4. **The arithmetic obstruction and determinant forms.** We say that a lattice $\Lambda < \mathrm{M}_d(\mathbb{R})$ is *determinant-rational* if

$$\det(\Lambda) \subset \lambda\mathbb{Q} \qquad \text{for some } \lambda \in \mathbb{R}^\times.$$

As shown in Theorem 2.6, this is equivalent to

$$\det(\Lambda) \subset \lambda\mathbb{Z} \qquad \text{for some } \lambda \in \mathbb{R}^\times.$$

Thus the hypothesis in Theorem 1.1 is precisely that Λ is not determinant-rational. See Section 2 for an algebraic description of all such lattices.

The same arithmetic obstruction has a dynamical interpretation. After rescaling, we may regard Λ as a unimodular lattice and hence as a point of

$$X = \mathrm{SL}_{d^2}(\mathbb{R})/\,\mathrm{SL}_{d^2}(\mathbb{Z}).$$

The left-right action

$$(g,h)\cdot v = gvh^\intercal, \qquad g,h \in \mathrm{SL}_d(\mathbb{R}),$$

induces the tensor-product representation $(g,h) \mapsto g \otimes h$ on $\mathrm{M}_d(\mathbb{R}) \simeq \mathbb{R}^d \otimes \mathbb{R}^d$. We denote its image by

$$H = \mathrm{SL}_d(\mathbb{R}) \otimes \mathrm{SL}_d(\mathbb{R}) < \mathrm{SL}_{d^2}(\mathbb{R}).$$

Ratner's orbit-closure theorem, together with the maximality of H, gives a closed-or-dense dichotomy. We show that the closed case is precisely the determinant-rational case. This gives the following qualitative counterpart of Theorem 1.1.

Proposition 1.3. *If $\Lambda < \mathrm{M}_d(\mathbb{R})$ is not determinant-rational, then*

$$\overline{\det(\Lambda)} = \mathbb{R}.$$

Thus determinant-rationality is exactly the obstruction to density of the determinant values.

A convenient coordinate realization is obtained as follows. Choose a $\mathbb{Z}$-basis $\mathcal{B} = \{v_{ij} : 1 \le i,j \le d\}$ of Λ and set

$$F_{\Lambda,\mathcal{B}}(x) := \det\left(\sum_{i,j} x_{ij}v_{ij}\right), \qquad x = (x_{ij}) \in \mathrm{M}_d(\mathbb{R}).$$

We call $F_{\Lambda,\mathcal{B}}$ a *determinant form.* In these coordinates, Theorem 1.1 becomes the following statement.

Theorem 1.4. *Let F be a determinant form on $\mathrm{M}_d(\mathbb{R})$ with algebraic coefficients, and suppose that F is not proportional to a polynomial with rational coefficients. Fix a norm invariant under a maximal compact subgroup of the stabilizer of F in $\mathrm{SL}_{d^2}(\mathbb{R})$. Then there exists $c_F = c(F, \|\cdot\|) > 0$ such that, for every $a < b$,*

$$\#\{x \in \mathrm{M}_d(\mathbb{Z}) : \ \|x\| < T,\ a < F(x) < b,\ F(x) \neq 0\} \sim c_F(b-a)T^{d(d-1)}.$$

The same conclusion holds for every Diophantine determinant form that is not proportional to a rational polynomial; see Theorem 3.10.

1.5. **Relation to the quantitative Oppenheim problem.** When $d = 2$,

$$\det \begin{pmatrix} x & y \\ z & w \end{pmatrix} = xw - yz$$

is a quadratic form of signature $(2, 2)$. Thus the determinant problem in degree two is a special case of the quantitative Oppenheim problem for indefinite quadratic forms. The corresponding counting statements are covered by the work of Eskin–Margulis–Mozes [14], following the higher-dimensional work of Dani–Margulis [10] and Eskin–Margulis–Mozes [13]; see also [21] for the signature $(2, 1)$ case. Theorems 1.1 and 1.2 may therefore be viewed as higher-degree analogues of the quantitative Oppenheim theorem.

1.6. **Comparison with the circle method.** Higher-degree forms in many variables are classically studied by the Hardy–Littlewood circle method. Birch's theorem and its variants give asymptotic formulas under hypotheses requiring the singular locus of the form to be sufficiently small relative to the number of variables and the degree. For a single homogeneous form P of degree k in n variables, the standard Birch condition is of the shape

$$n - \dim V_P^* > (k-1)2^k,$$

where $V_P^* := \{x : \nabla P(x) = 0\}$ is the singular locus of the hypersurface $P = 0$; see [3].

The determinant is far outside this range. It is a form of degree d in $n = d^2$ variables, but its singular locus is large:

$$V_{\det}^* = \{X \in \mathrm{M}_d(\mathbb{R}) : \operatorname{rank} X \leq d-2\}.$$

This determinantal variety has dimension $d^2 - 4$. Hence

$$n - \dim V_{\det}^* = 4,$$

whereas Birch's condition would require this quantity to be larger than $(d-1)2^d$. Thus the obstruction is not merely that the determinant hypersurface is singular; rather, the determinant has too few variables for its degree from the viewpoint of the standard circle method, and its singular locus has codimension only 4.

The present paper instead exploits the special geometry of the determinant: the left–right action of $\mathrm{SL}_d(\mathbb{R}) \times \mathrm{SL}_d(\mathbb{R})$, the rank stratification of the determinantal variety, and Diophantine separation of rational Plücker directions from the exceptional column- and row-isotropic summands.

1.7. **Main ideas and new difficulties for $d \geq 3$.** We now outline the proof, emphasizing the points where the determinant problem differs from the quadratic Oppenheim problem. Put

$$G = \mathrm{SL}_{d^2}(\mathbb{R}), \qquad X = \mathrm{SL}_{d^2}(\mathbb{R}) / \mathrm{SL}_{d^2}(\mathbb{Z}).$$

After identifying $\mathrm{M}_d(\mathbb{R})$ with $\mathbb{R}^{d^2}$ and normalizing covolume, a lattice $\Lambda < \mathrm{M}_d(\mathbb{R})$ defines a point of X. The left-right action of $\mathrm{SL}_d(\mathbb{R}) \times \mathrm{SL}_d(\mathbb{R})$ on $\mathrm{M}_d(\mathbb{R})$ induces the tensor-product representation; we denote its image by

$$H = \mathrm{SL}_d(\mathbb{R}) \otimes \mathrm{SL}_d(\mathbb{R}) < G.$$

Ratner's orbit-closure theorem gives a closed-or-dense dichotomy for H-orbits in X. The uniqueness, up to scale, of the H-invariant homogeneous polynomial of degree d identifies the closed case with rationality of the determinant form. Hence, for a non-determinant-rational lattice, the H-orbit is dense, and Shah's theorem [30] on expanding translates gives equidistribution for compactly supported functions on X.

To pass from orbit equidistribution to counting, we use Siegel transforms. For a compactly supported function f on $\mathrm{M}_d(\mathbb{R})$, the Siegel transform is

$$\widehat{f}(\Delta) := \sum_{0 \neq v \in \Delta} f(v), \qquad \Delta \in X.$$

Thus $\widehat{f}(\Delta)$ represents a lattice-point count when f approximates the indicator function of a region in $\mathrm{M}_d(\mathbb{R})$. The difficulty is that $\widehat{f}$ is unbounded on X: it becomes large in the cusp. To apply Shah's theorem to such unbounded functions, we need a uniform moment estimate with exponent strictly larger than one. In the determinant problem this is proved for a modified Siegel transform, in which the exact isotropic contribution is removed and counted separately.

The cusp is controlled by the Margulis α-function. If V is an r-dimensional Δ-rational subspace, its Plücker vector

$$\mathsf{w}_{\Delta,V} \in \wedge^r \mathrm{M}_d(\mathbb{R})$$

has norm equal to the covolume of $\Delta \cap V$ in V. Thus, up to harmless normalizations,

$$\alpha(\Delta) = \max_{1 \leq r \leq d^2-1} \sup_V \|\mathsf{w}_{\Delta,V}\|^{-1},$$

where V ranges over r-dimensional Δ-rational subspaces. This explains why all exterior powers $\wedge^r \mathrm{M}_d(\mathbb{R})$ enter the argument.

The principal diagonal ray is dictated by the singular-value geometry of determinant level sets at large norm. In a fixed determinant window $a < \det v < b$, matrices of large norm are close to the singular hypersurface $\{\det = 0\}$. In the main compact singular-value regime, $d-1$ singular values grow with the common scale T, while the remaining singular value compensates so that the determinant remains bounded. The left–right action of

$$b_t = \mathrm{diag}(e^{-t}, \ldots, e^{-t}, e^{(d-1)t}), \qquad a_t = b_t \otimes b_t \in H,$$

with $t \simeq \frac{1}{2} \log T$, is the determinant-preserving normalization of this configuration. Thus the principal ray a_t is the dynamical direction which produces the main term. The same ray also governs the height estimates. In the quadratic Oppenheim problem, the relevant one-parameter subgroup has a short weight structure on the exterior powers appearing in the height: only the weights $-1, 0, 1$ occur. For the determinant, the exterior powers of matrix space decompose into many irreducible H-summands. For $1 \le r \le d^2 - 1$, the skew Cauchy formula gives

$$\wedge^r \mathrm{M}_d(\mathbb{R}) \simeq \bigoplus_{\substack{\lambda \vdash r \\ \lambda \subseteq d \times d}} \mathsf{S}_\lambda(\mathbb{R}^d) \otimes \mathsf{S}_{\lambda^\intercal}(\mathbb{R}^d),$$

where $\lambda = (\lambda_1 \ge \cdots \ge \lambda_d \ge 0)$ ranges over partitions of r whose Young diagram lies in the $d \times d$ box, S_λ denotes the corresponding Schur module, and $\lambda^\intercal$ denotes the conjugate partition. On $\mathsf{S}_\lambda(\mathbb{R}^d)$, the b_t-weights are

$$-|\lambda| + di, \qquad 0 \le i \le \lambda_1.$$

Thus a single exterior power may contain many irreducible summands, and a single summand may contain many negative, zero, and positive weights.

There are two special summands in $\wedge^{kd} \mathrm{M}_d(\mathbb{R})$, for each $1 \le k \le d-1$:

$$\mathsf{S}_{(d^k)}(\mathbb{R}^d) \otimes \mathbf{1} \qquad \text{and} \qquad \mathbf{1} \otimes \mathsf{S}_{(d^k)}(\mathbb{R}^d).$$

Their decomposable vectors are precisely the Plücker vectors of column- and row-isotropic subspaces. We call these two summands *exceptional*, and call the remaining irreducible summands *nonexceptional*. A rational subspace in a critical degree kd is called *quasi-null* if its Plücker vector is polynomially close to one of the exceptional summands. Under the Diophantine condition, and for the parameters used in the proof, quasi-null rational subspaces are exactly the rational isotropic ones.

With this terminology, the main new difficulties are the following.

(1) *Higher-degree sublevel estimates.* Even on the nonexceptional summands, the local contraction estimates are much more delicate than in the quadratic case. Fix one Schur factor $\mathsf{S}_\lambda(\mathbb{R}^d)$, and write

$$\mathsf{S}_\lambda(\mathbb{R}^d) = \bigoplus_{i=0}^{\lambda_1} V_i, \qquad \mu_i = -|\lambda| + di,$$

for its b_t-weight decomposition. Let ℓ be the first index such that $\mu_\ell \ge 0$, and let p_i denote projection onto V_i. If $u_\xi = \exp(Y_\xi)$ belongs to the expanding horospherical subgroup, then the coordinate at this first nonnegative weight has the triangular form

$$p_\ell(u_\xi v) = \sum_{i=0}^{\ell} \frac{1}{(\ell - i)!} Y_\xi^{\ell-i} p_i(v).$$

In the quadratic Oppenheim problem the analogous coordinate is at most affine-linear in the horospherical parameters, and the corresponding sublevel estimates are elementary. For the determinant with $d \ge 3$, ℓ may be larger than 1, and the expression above is in general a higher-degree vector-valued polynomial in $\xi \in \mathbb{R}^{d-1}$. The order of vanishing at a zero can be greater than one and can vary from one zero to another.

The sublevel estimate therefore requires both algebraic and analytic inputs. We use the semistandard Young tableau basis of the Schur modules, together with the triangular action of the raising operators E_{di} to prove the necessary dimension bounds for the zero loci of the homogeneous initial forms. We then use the results of Cluckers–Miller [8, 9] twice: first to pass from these homogeneous initial-form estimates to the full triangular polynomial $p_\ell(u_\xi v)$, and second to make the resulting negative-moment estimates uniform over compact families of normalized vectors.

These sublevel estimates imply local contraction inequalities, which provide the local input behind the $L^{1+\theta}$-moment estimate for the modified height.

(2) *The uniform $L^{1+\theta}$-bound for the modified height.* The ordinary Margulis height is not sufficient for the determinant problem, because the exceptional summands have critical exponent exactly 1. Geometrically, the exceptional summands are generated by Plücker vectors of isotropic subspaces, and these subspaces lie in $\{\det = 0\}$. In the top critical dimension $d(d-1)$, they may contribute on the same scale as the nonsingular main term. Thus they should not be treated merely as cusp error.

We introduce a modified height which omits all rational subspaces contained in quasi-null subspaces:

$$\widehat{\alpha}_{\eta,M}(h;\Lambda) = \max_V \|h\mathsf{w}_{\Lambda,V}\|^{-1},$$

where the maximum is taken over all Λ-rational subspaces V that are not contained in a quasi-null critical subspace. The central technical estimate is the following uniform moment bound: for some $\theta > 0$,

$$\sup_{t\geq 0} \int_K \widehat{\alpha}_{\eta,M}(a_t k;\Lambda)^{1+\theta}\, dk < \infty,$$

where K is a maximal compact subgroup of H. This estimate is the key input that allows us to apply Shah's equidistribution theorem to modified Siegel transforms after truncation.

The proof of this bound follows the broad modified-height, avoidance, and iteration framework developed in [21]. In the determinant setting, however, each step requires new input because exceptional summands occur in every critical degree kd, $1 \leq k \leq d-1$.

(a) We construct an auxiliary modified height $\widetilde{\alpha}_{\eta,M}$ and prove a global Margulis inequality for it outside an explicit exceptional set, as given in Theorem 6.9:

$$\int_{B_N(1)} \widetilde{\alpha}_{\eta,M}(a_s nh;\Delta)\, dn \ll e^{-cs}\widetilde{\alpha}_{\eta,M}(h;\Delta)\log\big(\widehat{\alpha}_{\eta,M}(h;\Delta)+2\big) + e^{Cs},$$

provided (h, Δ) is outside the exceptional set. Here $B_N(1)$ is the unit ball in the expanding horospherical subgroup N for a_t. The logarithmic factor reflects the possible number of admissible competing short rational subspaces, and the additive term is absorbed later by the choice of step sizes in the iteration.
The main new point is that rational subspaces interact through intersections and sums. Since exceptional summands occur in every degree divisible by d, these intersections and sums may again have exceptional projections. The standard intersection-sum inequality of Eskin–Margulis–Mozes [13] controls covolumes but not these exceptional projections. We use the Mother Inequality of Benoist–Quint [2] to supply the missing control.

(b) We prove an avoidance estimate for the exceptional set. This set records the event that a rational Plücker vector which is not already quasi-null for the base lattice is moved very close to an exceptional summand. Since exact isotropic subspaces need not form a finite collection for $d \geq 3$, the low-dimensional quadratic avoidance argument does not apply directly. Instead, we use a quantitative non-divergence estimate [24] for a prescribed family of Plücker vectors, together with the Diophantine condition, to rule out the bad alternative.

(c) We combine the Margulis inequality and the avoidance estimate by an iteration argument. The logarithmic loss and the additive exponential error in the Margulis inequality force a careful choice of step sizes, but the iteration ultimately gives the uniform $L^{1+\theta}$-bound for $\widehat{\alpha}_{\eta,M}$. This bound permits the truncation of the modified Siegel transform, the application of Shah's theorem to the truncations, and the removal of the truncation by a tail estimate.

(3) *Infinitely many exact isotropic contributions.* The exact isotropic directions omitted from the modified height must then be counted. In the low-dimensional quadratic cases, once near-isotropic directions have been controlled, the exact exceptional subspaces contributing on the main scale form a finite collection. For the determinant with $d \geq 3$, this finiteness can fail even for irrational algebraic lattices. The lattice $\Lambda_{\sqrt{2}}$ in (1.6) already contains infinitely many rank-$d(d-1)$ submodules on which the determinant vanishes identically.

Since each such submodule may contribute on the $T^{d(d-1)}$-scale, one cannot remove a fixed finite exceptional list. Nor can one sum the individual contributions without proving convergence and controlling overlaps. We prove structural results showing that the exact rational isotropic subspaces can nevertheless be organized into finitely many families. On each family, the singular counting problem becomes the counting of bounded-rank matrices in a rational lattice. We prove the corresponding weighted asymptotics, estimate same-type and mixed column–row overlaps, and use the isotropic noncoincidence condition to exclude additional main-order coincidences. This gives the singular constant in (1.5).

(4) *From the principal-ray dynamics to norm balls.* The last difficulty is the passage from the dynamical statement along the principal ray a_t to the original norm-ball counting problem. The K_0-invariance of the norm allows the norm condition to be expressed in singular-value variables. After the common scale has been separated, the remaining singular-value ratios enter as parameters in the fiber kernels. On compact singular-value windows these parameters stay in a compact set, so the modified Siegel-transform limit along the principal ray applies uniformly. The unbalanced singular-value regions are controlled separately by weighted shell estimates. The same fiber integrals identify the constant $C_{\|\cdot\|}$ in (1.3) and the limiting functional in (1.5).

Although the results of this paper include the case $d = 2$, the level of difficulty changes substantially between $d = 2$ and $d \geq 3$. For $d = 2$, the relevant horospherical coordinates are affine-linear, there is only one critical exterior degree, and the exact rational isotropic subspaces form a finite collection. Thus many of the main technical difficulties of the present paper disappear. In the separate paper [22], we present this simplified argument in a self-contained form. It gives a new determinant-form proof of the $(2, 2)$ quantitative Oppenheim theorem of Eskin–Margulis–Mozes, and may serve as a model case for the method used here.

Organization of the paper. Section 2 classifies determinant-rational lattices, and proves the closed-or-dense orbit dichotomy. Section 3 develops the height formalism and the Schur-functor decomposition, introduces the Diophantine, quasi-null, and isotropic notions, and states the main technical results. Sections 4–8 establish uniform integrability: Section 4 proves the sublevel estimates, Section 5 derives the local representation estimates, Section 6 globalizes them to lattice heights, Section 7 proves avoidance, and Section 8 combines contraction and avoidance by iteration. Sections 9–10 analyze the singular directions. Section 9 describes the structure of rational isotropic subspaces, and Section 10 proves the bounded-rank estimates and the singular asymptotic. Sections 11–12 convert the dynamical estimates into counting: Section 11 establishes the relevant fiber-integral identities, and Section 12 proves equidistribution for the modified Siegel transforms and derives the main asymptotic formulas. Finally, Section 13 proves that algebraic lattices and algebraic determinant forms are Diophantine and gives examples with both positive and vanishing singular constant.

Acknowledgements. The authors would like to thank Yale University and the Korea Institute for Advanced Study for their hospitality; a substantial portion of this joint work was carried out during visits to these institutions. Part of this research was conducted while the authors were visiting the Simons

Laufer Mathematical Sciences Institute (SLMath), which is supported by the National Science Foundation under Grant No. DMS-2424139. W.K. is supported by a Korea Institute for Advanced Study individual grant no. HP101301.

2. Determinant-rational lattices and the orbit-closure dichotomy

In this section, we characterize determinant-rational lattices in $\mathrm{M}_d(\mathbb{R})$ and relate this arithmetic condition to the orbit closure of the left-right action of $\mathrm{SL}_d(\mathbb{R}) \times \mathrm{SL}_d(\mathbb{R})$ on the space of unimodular lattices.

Let $d \geq 2$ and put $N = d^2$. We identify $\mathrm{M}_d(\mathbb{R})$ with $\mathbb{R}^N$ by sending the elementary matrices E_{ij} to the standard basis. Under this identification, the determinant is a homogeneous polynomial of degree d on $\mathbb{R}^N$.

The group $\mathrm{GL}_d(\mathbb{R}) \times \mathrm{GL}_d(\mathbb{R})$ acts on $\mathrm{M}_d(\mathbb{R})$ by

$$(g_1, g_2) \cdot v = g_1 v g_2^{\intercal} \qquad (g_1, g_2 \in \mathrm{GL}_d(\mathbb{R}), v \in \mathrm{M}_d(\mathbb{R})). \tag{2.1}$$

Given a lattice $\Lambda < \mathrm{M}_d(\mathbb{R})$, choose a $\mathbb{Z}$-basis $\mathcal{B} = \{v_{ij}\}$ of Λ. We associate to Λ the determinant form

$$F_\Lambda(x) = \det\Big(\sum_{i,j} x_{ij} v_{ij}\Big), \qquad x = (x_{ij}) \in \mathrm{M}_d(\mathbb{R}).$$

This polynomial is well defined up to composition with an element of $\mathrm{GL}_N(\mathbb{Z})$. We say that Λ is *determinant-rational* if

$$\det(\Lambda) = F_\Lambda(\mathbb{Z}^N) \subset \lambda\mathbb{Q} \qquad \text{for some } \lambda \in \mathbb{R}^\times.$$

2.1. Determinant forms over $\mathbb{Q}$.

We begin with an algebraic description of rational forms of the determinant. In addition to the usual reduced-norm forms, one must also allow the outer forms arising from the transpose component of the determinant stabilizer.

A quadratic étale algebra over $\mathbb{Q}$ is either a quadratic field extension of $\mathbb{Q}$ or the split algebra $\mathbb{Q} \times \mathbb{Q}$. We denote its nontrivial $\mathbb{Q}$-automorphism by $z \mapsto \overline{z}$.

Definition 2.1 (Determinant algebras)**.** A *determinant algebra of degree d over* $\mathbb{Q}$ is a triple

$$(K, A, \tau),$$

of one of the following two types.

(1) $K/\mathbb{Q}$ is a quadratic field extension, A is a central simple K-algebra of degree d, and τ is a unitary involution on A, i.e. a $\mathbb{Q}$-linear anti-automorphism satisfying

$$\tau^2 = \mathrm{id}, \qquad \tau(zx) = \overline{z}\,\tau(x) \quad (z \in K,\ x \in A).$$

(2) $K = \mathbb{Q} \times \mathbb{Q}$, $A = A_0 \times A_0^{\mathrm{op}}$ for some central simple $\mathbb{Q}$-algebra A_0 of degree d, and

$$\tau(x, y) = (y, x).$$

We write
$$J(A,\tau) := \{x \in A : \tau(x) = x\},$$
which is a d^2-dimensional $\mathbb{Q}$-vector space. For $x \in J(A,\tau)$, define
$$N_{A,\tau}(x) := \begin{cases} \mathrm{Nrd}_{A/K}(x), & \text{in the case (1)}, \\ \mathrm{Nrd}_{A_0/\mathbb{Q}}(x_0), & \text{in the case (2), where } x = (x_0, x_0) \in A_0 \times A_0^{\mathrm{op}}. \end{cases}$$
Then $N_{A,\tau}$ is a homogeneous polynomial of degree d, defined over $\mathbb{Q}$, on $J(A,\tau)$.

The case (1) gives the outer determinant forms, while the case (2) gives the inner forms, namely the usual reduced-norm forms on central simple $\mathbb{Q}$-algebras.

We say that a determinant algebra (K, A, τ) is *split over* $\mathbb{R}$ if there is an $\mathbb{R}$-linear isomorphism $\iota : J(A,\tau) \otimes_{\mathbb{Q}} \mathbb{R} \xrightarrow{\sim} \mathrm{M}_d(\mathbb{R})$ such that $N_{A,\tau}(x) = \det(\iota(x))$. Only determinant algebras split over $\mathbb{R}$ will occur below. An order in A means a subring $\mathcal{O}$ which is a full $\mathbb{Z}$-lattice in the underlying $\mathbb{Q}$-vector space. If $\tau(\mathcal{O}) = \mathcal{O}$, we write $\mathcal{O}^\tau := \mathcal{O} \cap J(A,\tau)$, which is a full lattice in $J(A,\tau)$.

The following Frobenius–Dieudonné theorem (see [25, Theorem 2]) describes the linear maps preserving the determinant up to a scalar.

Lemma 2.2. *Let L be a field of characteristic zero, and let $T \in \mathrm{GL}(\mathrm{M}_d(L))$ satisfy*
$$\det(Tv) = c\,\det(v) \qquad (v \in \mathrm{M}_d(L))$$
for some $c \in L^\times$. Then either
$$Tv = g_1 v g_2 \qquad (v \in \mathrm{M}_d(L)),$$
or
$$Tv = g_1 v^\intercal g_2 \qquad (v \in \mathrm{M}_d(L)),$$
for some $g_1, g_2 \in \mathrm{GL}_d(L)$ satisfying $\det(g_1)\det(g_2) = c$.

We shall use the following elementary uniqueness statement.

Lemma 2.3. *If $g_1, g_2 \in \mathrm{GL}_d(L)$ satisfy*
$$g_1 v g_2 = v \qquad \textit{for all } v \in \mathrm{M}_d(L),$$
then there exists $c \in L^\times$ such that
$$g_1 = cI_d, \qquad g_2 = c^{-1} I_d.$$

Proof. Taking $v = I_d$, we get $g_1 g_2 = I_d$, so $g_2 = g_1^{-1}$. Hence $g_1 v g_1^{-1} = v$ for every $v \in \mathrm{M}_d(L)$. Therefore g_1 lies in the center of $\mathrm{M}_d(L)$, and so $g_1 = cI_d$ for some $c \in L^\times$. Then $g_2 = c^{-1} I_d$. □

We also record a simple rationality criterion for polynomials.

Lemma 2.4. *Let $p \in \mathbb{R}[x_1, \ldots, x_N]$. If $p(\mathbb{Z}^N) \subset \mathbb{Q}$, then $p \in \mathbb{Q}[x_1, \ldots, x_N]$.*

Proof. We first consider the case $N = 1$. Let $p \in \mathbb{R}[x]$ have degree m. Denote by
$$\Delta p(x) := p(x+1) - p(x)$$
the finite-difference operator. Then $\Delta^m p$ is constant and equal to $m!a_m$, where a_m is the leading coefficient of p. Moreover, each $\Delta^k p(0)$ is an integer linear combination of $p(0), \dots, p(k)$. Since $p(\mathbb{Z}) \subset \mathbb{Q}$, all these finite differences are rational. It follows that every coefficient of p is rational.

For general N, we argue by induction. Write
$$p(x_1, \dots, x_N) = \sum_{j=0}^{m} p_j(x_1, \dots, x_{N-1}) x_N^j.$$
For each integer r, the polynomial $p(\cdot, r)$ takes rational values on $\mathbb{Z}^{N-1}$. By the induction hypothesis, it has rational coefficients. Evaluating at $m+1$ distinct integers r and inverting the corresponding Vandermonde matrix, we conclude that each p_j has rational coefficients. □

We use the following standard classification of twisted forms of the determinant [31].

Lemma 2.5 (Twisted determinant forms)**.** *Let $F \in \mathbb{Q}[x_1, \dots, x_N]$ be a homogeneous polynomial of degree d. Suppose that F becomes linearly equivalent to* det *over $\overline{\mathbb{Q}}$. Then there exist a determinant algebra (K, A, τ) of degree d over $\mathbb{Q}$ and a $\mathbb{Q}$-basis $\{w_{ij}\}$ of $J(A, \tau)$ such that*
$$F(x) = c\, N_{A,\tau}\left(\sum_{i,j} x_{ij} w_{ij}\right)$$
for some $c \in \mathbb{Q}^\times$. Conversely, every polynomial of this form becomes linearly equivalent to the determinant over $\overline{\mathbb{Q}}$.

Moreover, F is linearly equivalent over $\mathbb{R}$ to the ordinary determinant if and only if (K, A, τ) is split over $\mathbb{R}$.

Proof. This is the classification of twisted forms of the determinant, due to Waterhouse; see [31]. We recall the descent mechanism to fix the notation.

Choose $g \in \mathrm{GL}_N(\overline{\mathbb{Q}})$ such that
$$F = \det \circ g.$$
For each $\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, the element $z_\sigma := g\, \sigma(g)^{-1}$ lies in the determinant stabilizer. By Lemma 2.2, it belongs either to the left-right component or to the transpose-left-right component. Projection to the component group gives a quadratic character $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathbb{Z}/2\mathbb{Z}$. This character determines the quadratic étale algebra $K/\mathbb{Q}$. Over K, the cocycle lands in the identity component of the determinant stabilizer and hence gives, by the usual descent correspondence for matrix algebras, a central simple K-algebra A of degree d. The action of the nontrivial automorphism of $K/\mathbb{Q}$, together with the transpose component, descends to a unitary involution τ on A. The descended form is precisely the restriction of $\mathrm{Nrd}_{A/K}$ to $J(A, \tau)$.

Conversely, after extension of scalars to $\overline{\mathbb{Q}}$, every determinant algebra becomes the split matrix algebra with its standard transpose descent datum, and $N_{A,\tau}$ becomes the ordinary determinant. The last assertion follows by the same argument after extending scalars to $\mathbb{R}$. □

We shall also use the following terminology. A determinant form F over $\mathbb{Q}$ is called $\mathbb{Q}$-*split* if it is proportional to $\det \circ T$ for some $T \in \mathrm{GL}_N(\mathbb{Q})$. We say that a lattice Λ is of $\mathbb{Q}$-split type if there exist $g_1, g_2 \in \mathrm{GL}_d(\mathbb{R})$ such that

$$\Lambda \quad \text{is commensurable with} \quad g_1 \, \mathrm{M}_d(\mathbb{Z}) g_2.$$

We now give the arithmetic characterizations of determinant-rational lattices.

Theorem 2.6. *Let $\Lambda < \mathrm{M}_d(\mathbb{R})$ be a lattice. The following conditions are equivalent:*

(1) Λ is determinant-rational;
(2) $\det(\Lambda) \subset \lambda\mathbb{Z}$ for some $\lambda \in \mathbb{R}^\times$;
(3) F_Λ is proportional to a polynomial with rational coefficients;
(4) there exist a determinant algebra (K, A, τ) of degree d over $\mathbb{Q}$, split over $\mathbb{R}$, a τ-stable order $\mathcal{O} \subset A$, an $\mathbb{R}$-linear norm-form isomorphism

$$\iota : J(A, \tau) \otimes_{\mathbb{Q}} \mathbb{R} \xrightarrow{\sim} \mathrm{M}_d(\mathbb{R}),$$

and $g_1, g_2 \in \mathrm{GL}_d(\mathbb{R})$, such that Λ is commensurable with

$$g_1 \, \iota(\mathcal{O}^\tau) \, g_2;$$

(5) there exist a determinant algebra (K, A, τ) of degree d over $\mathbb{Q}$, split over $\mathbb{R}$, and a $\mathbb{Q}$-basis $\{w_{ij}\}$ of $J(A, \tau)$, such that

$$F_\Lambda(x) = c \, N_{A,\tau}\left(\sum_{i,j} x_{ij} w_{ij} \right) \quad \text{for some } c \in \mathbb{R}^\times.$$

Moreover, Λ is of $\mathbb{Q}$-split type if and only if F_Λ is $\mathbb{Q}$-split.

Proof. The implication (2) ⇒ (1) is immediate. If (3) holds, then there are $c \in \mathbb{R}^\times$ and $P \in \mathbb{Q}[x_1, \ldots, x_N]$ such that $F_\Lambda = cP$. Since the values of P on $\mathbb{Z}^N$ have bounded denominators, there exists $m \geq 1$ such that $mP(\mathbb{Z}^N) \subset \mathbb{Z}$. Thus $\det(\Lambda) = F_\Lambda(\mathbb{Z}^N) \subset c\, m^{-1}\mathbb{Z}$, which proves (3) ⇒ (2).

We next prove (1) ⇒ (3). If Λ is determinant-rational, then $F_\Lambda(\mathbb{Z}^N) \subset \lambda\mathbb{Q}$ for some $\lambda \in \mathbb{R}^\times$. Hence $\lambda^{-1} F_\Lambda$ takes rational values on $\mathbb{Z}^N$. By Lemma 2.4,

$$\lambda^{-1} F_\Lambda \in \mathbb{Q}[x_1, \ldots, x_N],$$

so F_Λ is proportional to a polynomial with rational coefficients.

We now prove (3) ⇒ (5). After rescaling, we may assume that $F_\Lambda \in \mathbb{Q}[x_1, \ldots, x_N]$. Since F_Λ is a real determinant form, it is linearly equivalent to det over $\mathbb{R}$, and hence over $\overline{\mathbb{Q}}$. By Lemma 2.5, there exist a determinant algebra (K, A, τ) over $\mathbb{Q}$ and a $\mathbb{Q}$-basis $\{w_{ij}\}$ of $J(A, \tau)$ such that

$$F_\Lambda(x) = c \, N_{A,\tau}\left(\sum_{i,j} x_{ij} w_{ij} \right)$$

for some $c \in \mathbb{R}^\times$. Since F_Λ is linearly equivalent over $\mathbb{R}$ to the ordinary determinant, the determinant algebra is split over $\mathbb{R}$.

We prove (5) $\Rightarrow$ (4). Let $\{v_{ij}\}$ be the $\mathbb{Z}$-basis of Λ used to define F_Λ, and let W be the $\mathbb{Z}$-span of the basis $\{w_{ij}\}$ in $J(A,\tau)$. Choose a τ-stable order $\mathcal{O} \subset A$. Then W and $\mathcal{O}^\tau$ are full lattices in the same $\mathbb{Q}$-vector space $J(A,\tau)$, hence are commensurable.

Let $L : J(A,\tau) \otimes_{\mathbb{Q}} \mathbb{R} \to \mathrm{M}_d(\mathbb{R})$ be the real linear isomorphism determined by $L(w_{ij}) = v_{ij}$. Let $\iota : J(A,\tau) \otimes_{\mathbb{Q}} \mathbb{R} \xrightarrow{\sim} \mathrm{M}_d(\mathbb{R})$ be a norm-form isomorphism, so that

$$N_{A,\tau}(x) = \det(\iota(x)).$$

Then $T := L \circ \iota^{-1}$ satisfies

$$\det(Tv) = c \det(v) \qquad (v \in \mathrm{M}_d(\mathbb{R})).$$

By Lemma 2.2, either

$$T(v) = g_1 v g_2 \quad \text{or} \quad T(v) = g_1 v^\intercal g_2$$

for some $g_1, g_2 \in \mathrm{GL}_d(\mathbb{R})$. In the second case, replace ι by the norm-form isomorphism $x \mapsto \iota(x)^\intercal$. Thus, after this harmless change of ι, we have $L(x) = g_1 \iota(x) g_2$. Since $L(W) = \Lambda$ and W is commensurable with $\mathcal{O}^\tau$, it follows that Λ is commensurable with $g_1\, \iota(\mathcal{O}^\tau)\, g_2$. This proves (4).

Finally, we prove (4) $\Rightarrow$ (2). Let $x \in \mathcal{O}^\tau$. Since $\mathcal{O}$ is an order, $\mathrm{Nrd}_{A/K}(x)$ is an algebraic integer of K. Since $x = \tau(x)$, this reduced norm is fixed by the nontrivial automorphism of $K/\mathbb{Q}$, and hence lies in $\mathbb{Z}$. Therefore

$$\det(g_1 \iota(x) g_2) = \det(g_1 g_2)\, N_{A,\tau}(x) \in \det(g_1 g_2)\mathbb{Z}.$$

If Λ is commensurable with $g_1 \iota(\mathcal{O}^\tau) g_2$, then there exists $m \geq 1$ such that $m\Lambda \subset g_1 \iota(\mathcal{O}^\tau) g_2$. Consequently,

$$\det(\Lambda) \subset m^{-d} \det(g_1 g_2)\mathbb{Z}.$$

This proves (2). It remains to prove the final assertion. Let

$$L_\Lambda : \mathrm{M}_d(\mathbb{R}) \to \mathrm{M}_d(\mathbb{R})$$

be the linear isomorphism determined by the chosen $\mathbb{Z}$-basis $\{v_{ij}\}$ of Λ, so that $L_\Lambda(E_{ij}) = v_{ij}$. Then

$$\Lambda = L_\Lambda\, \mathrm{M}_d(\mathbb{Z}) \qquad \text{and} \qquad F_\Lambda = \det \circ L_\Lambda.$$

Recall that, for $S_1, S_2 \in \mathrm{GL}_N(\mathbb{R})$, the lattices $S_1\mathbb{Z}^N$ and $S_2\mathbb{Z}^N$ are commensurable if and only if $S_1^{-1} S_2 \in \mathrm{GL}_N(\mathbb{Q})$.

Suppose first that Λ is of $\mathbb{Q}$-split type. Then there exist $g_1, g_2 \in \mathrm{GL}_d(\mathbb{R})$ such that Λ is commensurable with $g_1\, \mathrm{M}_d(\mathbb{Z}) g_2$. Let

$$S : \mathrm{M}_d(\mathbb{R}) \to \mathrm{M}_d(\mathbb{R}), \qquad S(X) = g_1 X g_2.$$

The lattices $L_\Lambda\, \mathrm{M}_d(\mathbb{Z})$ and $S\, \mathrm{M}_d(\mathbb{Z})$ are commensurable. Hence

$$T := S^{-1} L_\Lambda \in \mathrm{GL}_N(\mathbb{Q}).$$

It follows that

$$F_\Lambda(X) = \det(L_\Lambda X) = \det(S(TX)) = \det(g_1)\det(g_2)\det(TX).$$

Thus

$$F_\Lambda = \det(g_1)\det(g_2)\ \det \circ T,$$

and therefore F_Λ is $\mathbb{Q}$-split.

Conversely, suppose that F_Λ is $\mathbb{Q}$-split. Then there exist $c \in \mathbb{R}^\times$ and $T \in \mathrm{GL}_N(\mathbb{Q})$ such that $F_\Lambda = c \det \circ T$. Set $S := L_\Lambda T^{-1}$. For every $X \in \mathrm{M}_d(\mathbb{R})$, we have

$$\det(SX) = \det(L_\Lambda T^{-1} X) = c \det(X).$$

By Lemma 2.2, there exist $g_1, g_2 \in \mathrm{GL}_d(\mathbb{R})$ such that either $S(X) = g_1 X g_2$ for all X, or $S(X) = g_1 X^\intercal g_2$ for all X. Since $T \in \mathrm{GL}_N(\mathbb{Q})$, the lattice $T\,\mathrm{M}_d(\mathbb{Z})$ is commensurable with $\mathrm{M}_d(\mathbb{Z})$. Hence

$$\Lambda = L_\Lambda \mathrm{M}_d(\mathbb{Z}) = S\big(T\,\mathrm{M}_d(\mathbb{Z})\big)$$

is commensurable with $S\,\mathrm{M}_d(\mathbb{Z}) = g_1\,\mathrm{M}_d(\mathbb{Z}) g_2$. Thus Λ is of $\mathbb{Q}$-split type. □

2.2. **Orbit closures.** Let

$$G = \mathrm{SL}_N(\mathbb{R}), \qquad \Gamma = \mathrm{SL}_N(\mathbb{Z}), \qquad X = G/\Gamma.$$

The space X parametrizes unimodular lattices in $\mathrm{M}_d(\mathbb{R})$. For $g = (g_{ij}) \in \mathrm{SL}_d(\mathbb{R})$ and $h \in \mathrm{SL}_d(\mathbb{R})$, their Kronecker product is

$$g \otimes h = \begin{pmatrix} g_{11}h & \cdots & g_{1d}h \\ \vdots & \ddots & \vdots \\ g_{d1}h & \cdots & g_{dd}h \end{pmatrix} \in \mathrm{SL}_N(\mathbb{R}).$$

We set

$$H = \mathrm{SL}_d(\mathbb{R}) \otimes \mathrm{SL}_d(\mathbb{R}).$$

Under the identification $E_{ij} = e_i \otimes e_j$, one has

$$(h_1 \otimes h_2)v = h_1 v h_2^\intercal.$$

The following consequence of Ratner's orbit-closure theorem is the dynamical input.

Theorem 2.7. *For every $[g] \in G/\Gamma$, the orbit $H[g]$ is either closed or dense. If $H[g]$ is closed, then $H \cap g\Gamma g^{-1}$ is a lattice in H.*

Proof. The group H is connected, semisimple, has no compact factors, and is generated by one-parameter unipotent subgroups. Moreover, H is a maximal connected Lie subgroup of G; see [12]. Ratner's orbit-closure theorem [28] therefore implies that the closure of $H[g]$ is either $H[g]$ or all of G/Γ. In the closed case, the stabilizer $H \cap g\Gamma g^{-1}$ is a lattice in H. □

We next identify the invariant polynomials for the H-action.

Lemma 2.8. *For any $g \in G$, the space of homogeneous degree-d polynomials on $\mathrm{M}_d(\mathbb{R})$ invariant under $g^{-1}Hg$ is one-dimensional and is spanned by*

$$F_g := \det \circ g.$$

Proof. Let f be a homogeneous degree-d polynomial invariant under $g^{-1}Hg$ for some $g \in G$. After replacing f by $f \circ g^{-1}$, it suffices to treat the case $g = e$. If $f(I_d) = 0$, then f vanishes on the H-orbit of I_d, which is $\mathrm{SL}_d(\mathbb{R})$. By homogeneity, f vanishes on the positive-determinant open subset $\mathbb{R}^+ \mathrm{SL}_d(\mathbb{R})$, which is Zariski dense in $\mathrm{M}_d(\mathbb{R})$. Hence it vanishes identically as a polynomial. Thus we may assume $f(I_d) = 1$.

For $X \in \mathrm{GL}_d(\mathbb{R})$, the H-orbit of X is determined by $\det X$. Hence

$$f(X) = f\big(\mathrm{diag}(\det X, 1, \dots, 1)\big).$$

For every $t \neq 0$, $\mathrm{diag}(t^d, 1, \dots, 1)$ lies in the H-orbit of tI_d. Therefore

$$f\big(\mathrm{diag}(t^d, 1, \dots, 1)\big) = f(tI_d) = t^d.$$

It follows that for all $s > 0$, $f\big(\mathrm{diag}(s, 1, \dots, 1)\big) = s$. Hence $f(X) = \det X$ on the open set $\{X : \det X > 0\}$. Since this set is Zariski dense in $\mathrm{M}_d(\mathbb{R})$, the two polynomials agree identically. □

We can now characterize determinant-rational lattices dynamically.

Theorem 2.9. *Let $g \in G$, and let $\Lambda = g\mathbb{Z}^N$. Then:*

(1) Λ is determinant-rational if and only if $H[g]$ is closed in X.

(2) If Λ is not determinant-rational, then $H[g]$ is dense in X, and

$$\overline{\det(\Lambda)} = \mathbb{R}.$$

Proof. We first prove that closedness implies determinant-rationality. Suppose that $H[g]$ is closed. By Theorem 2.7, $\Gamma \cap g^{-1}Hg$ is a lattice in $g^{-1}Hg$.

Let W_0 be the space of homogeneous degree-d polynomials on $\mathbb{R}^N$ invariant under $\Gamma \cap g^{-1}Hg$. This is a $\mathbb{Q}$-subspace of $\mathrm{Sym}^d((\mathbb{R}^N)^*)$, since $\Gamma \cap g^{-1}Hg \subset \mathrm{SL}_N(\mathbb{Z})$. Moreover $F_g \in W_0$. By the Borel density theorem, $\Gamma \cap g^{-1}Hg$ is Zariski dense in $g^{-1}Hg$; see, for instance, [32]. Hence every element of W_0 is invariant under $g^{-1}Hg$. By Lemma 2.8,

$$W_0 = \mathbb{R}F_g.$$

Since W_0 is defined over $\mathbb{Q}$, the line $\mathbb{R}F_g$ is defined over $\mathbb{Q}$. Hence F_g, and therefore F_Λ, is proportional to a polynomial with rational coefficients. By Theorem 2.6, Λ is determinant-rational.

Conversely, suppose that $\Lambda = g\mathbb{Z}^N$ is determinant-rational. After rescaling F_g, we may assume that $F_g \in \mathbb{Q}[x_1, \dots, x_N]$. Let

$$\mathbf{L} := \mathrm{Stab}_{\mathrm{SL}_N}(F_g)^\circ.$$

Since F_g has rational coefficients, $\mathbf{L}$ is an algebraic group defined over $\mathbb{Q}$. Moreover, $\mathbf{L}(\mathbb{R})^\circ = g^{-1}Hg$. The group $\mathbf{L}$ is semisimple. By the Borel–Harish-Chandra theorem [5], $\mathbf{L}(\mathbb{Q}) \cap \mathrm{SL}_N(\mathbb{Z})$ is a lattice in $\mathbf{L}(\mathbb{R})^\circ$. Hence $g^{-1}Hg \cap \Gamma$ is a lattice in $g^{-1}Hg$. Therefore $H[g]$ is closed.

This proves (1). Part (2) follows immediately from Theorem 2.7: if Λ is not determinant-rational, then $H[g]$ is not closed, and hence is dense in X. Since $\Lambda = g\mathbb{Z}^N \supset g\Gamma e_1$ and $Ge_1 = \mathbb{R}^N - \{0\}$, we obtain

$$\overline{\det(\Lambda)} \supset \det\big(\overline{Hg\Gamma}\, e_1\big) = \det(Ge_1) = \det(\mathrm{M}_d(\mathbb{R}) - \{0\}) = \mathbb{R}.$$

Thus $\overline{\det(\Lambda)} = \mathbb{R}$. □

3. Height functions and main technical results

In this section we introduce the height functions and the representation-theoretic framework used in the proof of Theorem 3.8. We then define the Diophantine and quasi-null conditions that isolate the exceptional subspaces for the action of $H = \mathrm{SL}_d(\mathbb{R}) \times \mathrm{SL}_d(\mathbb{R})$, and state the principal technical results proved in the subsequent sections.

3.1. **Rational subspaces of a lattice.** Put $N = d^2$, and identify $\mathrm{M}_d(\mathbb{R})$ with $\mathbb{R}^N$ by the coordinate map $(x_{ij})_{1\le i,j\le d} \to \sum_{i,j} x_{ij}E_{ij}$ where E_{ij} is the elementary matrix whose only nonzero entry is one at the (i,j) position. We equip $\mathrm{M}_d(\mathbb{R})$ with the Euclidean inner product

$$\langle v, w\rangle = \mathrm{tr}(vw^{\mathsf{T}}),$$

and use the induced inner products on all exterior powers.

Let $V < \mathrm{M}_d(\mathbb{R})$ be an r-dimensional subspace for $1 \le r \le d^2 - 1$. A *Plücker vector* of V is a nonzero decomposable vector

$$\mathsf{w}_V = v_1 \wedge \cdots \wedge v_r \in \wedge^r \mathrm{M}_d(\mathbb{R}), \tag{3.1}$$

where $v_1, \ldots, v_r$ is a basis of V. Thus the Plücker line $\mathbb{R}\mathsf{w}_V$ depends only on V.

Let $\Delta < \mathrm{M}_d(\mathbb{R})$ be a lattice. A subspace $V < \mathrm{M}_d(\mathbb{R})$ is Δ-*rational* if $\Delta \cap V$ is a lattice in V. For such a subspace, its Δ-*Plücker vector* is

$$\mathsf{w}_{\Delta,V} = v_1 \wedge \cdots \wedge v_r, \tag{3.2}$$

where $v_1, \ldots, v_r$ is a $\mathbb{Z}$-basis of $\Delta \cap V$. This vector is well-defined up to sign, and

$$\|\mathsf{w}_{\Delta,V}\| = \mathrm{covol}_V(\Delta \cap V).$$

We write

$$d_\Delta(V) := \begin{cases} 1, & V = \{0\}, \\ \|\mathsf{w}_{\Delta,V}\|, & V \ne \{0\}. \end{cases}$$

3.2. **The height function α.** Let

$$X = \mathrm{SL}_N(\mathbb{R})/\,\mathrm{SL}_N(\mathbb{Z})$$

be the space of unimodular lattices in $\mathrm{M}_d(\mathbb{R})$. The α-*characteristics*, introduced by Eskin–Margulis–Mozes in [13], are the functions $\alpha_1, \ldots, \alpha_{N-1} : X \to (0,\infty)$ defined by

$$\alpha_i(\Delta) := \sup\left\{\frac{1}{d_\Delta(V)} : V < \mathrm{M}_d(\mathbb{R}) \text{ is } \Delta\text{-rational and } \dim V = i\right\}.$$

Equivalently, $\alpha_i(\Delta)^{-1}$ is the least covolume of an i-dimensional primitive sublattice of Δ. The usual height is

$$\alpha(\Delta) := \max_{1\le i\le N-1} \alpha_i(\Delta).$$

3.3. **Schur functors.** Let

$$H_1 = \mathrm{SL}_d(\mathbb{R}) \times \{e\}, \qquad H_2 = \{e\} \times \mathrm{SL}_d(\mathbb{R}), \qquad H = H_1 H_2.$$

The action

$$(g,h)\cdot X = gvh^{\mathsf{T}} \qquad (g,h \in \mathrm{SL}_d(\mathbb{R}),\ v \in \mathrm{M}_d(\mathbb{R}))$$

induces an action of H on every exterior power of $\mathrm{M}_d(\mathbb{R})$.

Let C and R be two copies of $\mathbb{R}^d$, representing the column and row coordinates, respectively. Under the identification $E_{ij} \leftrightarrow e_i \otimes f_j$, one has

$$\mathrm{M}_d(\mathbb{R}) \simeq \mathsf{C} \otimes \mathsf{R},$$

and the two factors of H act in the standard way on C and R.

The exterior powers of matrix space decompose into irreducible H-summands by the skew Cauchy formula, that is, the Cauchy formula for exterior powers:

$$\wedge^r \mathrm{M}_d(\mathbb{R}) \simeq \bigoplus_{\substack{\lambda\vdash r \\ \lambda\subseteq d\times d}} \mathsf{S}_\lambda(\mathbb{R}^d) \otimes \mathsf{S}_{\lambda^{\mathsf{T}}}(\mathbb{R}^d),$$

where $\lambda = (\lambda_1 \ge \cdots \ge \lambda_d \ge 0)$ ranges over partitions of r whose Young diagram lies in the $d \times d$ box, S_λ denotes the corresponding Schur module, and λ^{T} denotes the conjugate partition. We use the standard notation for Schur functors; see [17, Section 4.2] or [16, Chapter I].

If $\lambda = (\lambda_1, \dots, \lambda_d)$, with $\lambda_1 \ge \cdots \ge \lambda_d \ge 0$, then the restriction of $\mathsf{S}_\lambda(\mathbb{R}^d)$ to $\mathrm{SL}_d(\mathbb{R})$ has highest weight

$$\sum_{j=1}^{d-1} (\lambda_j - \lambda_{j+1})\omega_j.$$

It is therefore trivial precisely when $\lambda = (q^d)$ for some $q \ge 0$. Consequently, a trivial factor can occur in $\wedge^i \mathrm{M}_d(\mathbb{R})$ only when $d \mid i$.

For $1 \le k \le d-1$, define

$$\mathcal{M}_{k,1} := \left(\wedge^{kd} \mathrm{M}_d(\mathbb{R})\right)^{H_2} \simeq \mathsf{S}_{(d^k)}(\mathsf{C}) \otimes \mathbf{1},$$

and

$$\mathcal{M}_{k,2} := \left(\wedge^{kd} \mathrm{M}_d(\mathbb{R})\right)^{H_1} \simeq \mathbf{1} \otimes \mathsf{S}_{(d^k)}(\mathsf{R}).$$

Thus $\mathcal{M}_{k,1}$ and $\mathcal{M}_{k,2}$ are the maximal subspaces on which H_2 and H_1, respectively, act trivially. Since the skew Cauchy decomposition is multiplicity-free, it gives an H-equivariant direct-sum decomposition

$$\wedge^{kd} \mathrm{M}_d(\mathbb{R}) = \mathcal{M}_{k,0} \oplus \mathcal{M}_{k,1} \oplus \mathcal{M}_{k,2}, \tag{3.3}$$

where $\mathcal{M}_{k,0}$ is the sum of the remaining irreducible summands. Set

$$\mathcal{M}_m := \bigoplus_{k=1}^{d-1} \mathcal{M}_{k,m}, \qquad 0 \le m \le 2,$$

and let

$$\pi_{k,m} : \wedge^{kd} \mathrm{M}_d(\mathbb{R}) \to \mathcal{M}_{k,m}$$

be the H-equivariant projection associated with (3.3). For $r \ge 1$ and $m \in \{1,2\}$, put

$$\pi_{k,m}^{(r)} := \wedge^r \pi_{k,m} : \wedge^r\big(\wedge^{kd} \mathrm{M}_d(\mathbb{R})\big) \to \wedge^r \mathcal{M}_{k,m}.$$

Thus

$$\pi_{k,m}^{(r)}(w_1 \wedge \cdots \wedge w_r) = \pi_{k,m}(w_1) \wedge \cdots \wedge \pi_{k,m}(w_r).$$

3.4. **Isotropic subspaces.** For a subspace $U < \mathbb{R}^d$, define

$$\mathcal{L}(U) := \{X \in \mathrm{M}_d(\mathbb{R}) : \mathrm{Col}(X) \subset U\} \text{ and } \mathcal{R}(U) := \{X \in \mathrm{M}_d(\mathbb{R}) : \mathrm{Row}(X) \subset U\}.$$

Both spaces have dimension $d \dim U$.

Definition 3.1. Let $1 \le k \le d-1$. A kd-dimensional subspace $V < \mathrm{M}_d(\mathbb{R})$ is *column-isotropic* if $V = \mathcal{L}(U)$ for some k-dimensional subspace $U < \mathbb{R}^d$, and is *row-isotropic* if $V = \mathcal{R}(U)$ for such a subspace U. We call V *isotropic* if it is column-isotropic or row-isotropic.

Every isotropic subspace is contained in the singular hypersurface $\{X : \det X = 0\}$. We use the term *isotropic* only for the two families above; an arbitrary linear subspace on which the determinant vanishes identically need not be isotropic in this sense.

The exceptional summands in (3.3) have the following geometric interpretation.

Lemma 3.2. *Let $V < \mathrm{M}_d(\mathbb{R})$ have dimension kd, and let w_V be a Plücker vector of V. Then*

$$V \text{ is column-isotropic} \iff \mathsf{w}_V \in \mathcal{M}_{k,1};$$

$$V \text{ is row-isotropic} \iff \mathsf{w}_V \in \mathcal{M}_{k,2}.$$

Proof. We prove the column statement; the row statement follows by interchanging the two factors of H.

Suppose first that $V = \mathcal{L}(U)$ with $\dim U = k$. Right multiplication preserves column spaces, so H_2 preserves V. Under the identification $V \simeq U \otimes \mathsf{R}$, the determinant of the action of $h \in H_2$ on V is $(\det h)^k = 1$. Hence H_2 fixes w_V, and therefore $\mathsf{w}_V \in \mathcal{M}_{k,1}$.

Conversely, suppose that $\mathsf{w}_V \in \mathcal{M}_{k,1}$. Then w_V is fixed by H_2. Since a nonzero decomposable exterior vector determines its underlying subspace, V is H_2-invariant. As an H_2-module,

$$\mathrm{M}_d(\mathbb{R}) \simeq \mathsf{C} \otimes \mathsf{R}$$

is the tensor product of a trivial multiplicity space C with the irreducible standard H_2-module R. Hence every H_2-submodule is of the form $U \otimes \mathsf{R}$ for a unique subspace $U < \mathsf{C}$. Since $\dim V = kd$, one has $\dim U = k$, and thus $V = \mathcal{L}(U)$. ☐

3.5. **A rationality obstruction for exceptional summands.** The next observation clarifies the relation between determinant-rationality and the exceptional summands. A non-determinant-rational lattice may contain infinitely many rational isotropic subspaces, as in the example $\Lambda_{\sqrt{2}}$ from the introduction. What cannot happen is that the Plücker vectors of such subspaces rationally fill an entire exceptional summand.

Proposition 3.3 (A rational exceptional summand forces determinant-rationality)**.** *Let $\Delta < \mathrm{M}_d(\mathbb{R})$ be a lattice and $1 \leq k \leq d-1$. Suppose that $\mathcal{M}_{k,1} \subset \wedge^{kd} \mathrm{M}_d(\mathbb{R})$ is rational with respect to the exterior lattice $\wedge^{kd}\Delta$. Then Δ is determinant-rational. The analogous statement holds with $\mathcal{M}_{k,2}$ in place of $\mathcal{M}_{k,1}$.*

Proof. We prove the column case; the row case is obtained by interchanging the two tensor factors. Let $\Delta = g\mathbb{Z}^N$. Put
$$M_g := (\wedge^{kd} g^{-1})\mathcal{M}_{k,1} \subset \wedge^{kd}\mathbb{R}^N.$$
By assumption M_g is a $\mathbb{Q}$-rational subspace of $\wedge^{kd}\mathbb{R}^N$. Hence its pointwise stabilizer
$$\mathbf{L} := \Big\{h \in \mathrm{SL}_N : \ (\wedge^{kd} h)v = v \text{ for every } v \in M_g\Big\}$$
is a $\mathbb{Q}$-algebraic subgroup of SL_N.

We claim that
$$\mathbf{L}(\mathbb{R})^\circ = g^{-1} H_2 g. \tag{3.4}$$
It is enough to prove that the identity component of the pointwise stabilizer of $\mathcal{M}_{k,1}$ in $\mathrm{SL}(\mathsf{C} \otimes \mathsf{R})$ is H_2.

For $U \in \mathrm{Gr}(k, \mathsf{C})$, set
$$W_U := U \otimes \mathsf{R}.$$
Let P denote the pointwise stabilizer of $\mathcal{M}_{k,1}$ in $\mathrm{SL}(\mathsf{C}\otimes\mathsf{R})$. Let $Z \in \mathfrak{sl}(\mathsf{C}\otimes\mathsf{R})$ belong to $\mathrm{Lie}(P)$. Then
$$Z\mathsf{w}_{W_U} = 0 \qquad \text{for every } U \in \mathrm{Gr}(k, \mathsf{C}).$$
This implies that
$$Z(U \otimes \mathsf{R}) \subset U \otimes \mathsf{R} \qquad \text{for every } U \in \mathrm{Gr}(k, \mathsf{C}).$$
We claim that Z preserves each subspace $\mathbb{R}c \otimes \mathsf{R}$, $0 \neq c \in \mathsf{C}$. If $k = 1$, this is immediate. If $k > 1$, then $\mathbb{R}c = \bigcap_{\substack{U \in \mathrm{Gr}(k,\mathsf{C}) \\ c \in U}} U$, and therefore $\mathbb{R}c \otimes \mathsf{R} = \bigcap_{\substack{U \in \mathrm{Gr}(k,\mathsf{C}) \\ c \in U}} (U \otimes \mathsf{R})$. Since Z preserves every $U \otimes \mathsf{R}$, it preserves this intersection.

Consequently, for every $0 \neq c \in \mathsf{C}$, there exists $B_c \in \mathrm{End}(\mathsf{R})$ such that
$$Z(c \otimes r) = c \otimes B_c r \qquad (r \in \mathsf{R}).$$

If c_1, c_2 are linearly independent, then applying the same statement to $c_1 + c_2$ and using linearity gives

$$c_1 \otimes B_{c_1} r + c_2 \otimes B_{c_2} r = (c_1 + c_2) \otimes B_{c_1+c_2} r \qquad (r \in \mathsf{R}).$$

Hence

$$B_{c_1} = B_{c_2} = B_{c_1+c_2}.$$

The same conclusion for proportional nonzero vectors follows by scaling. Thus B_c is independent of c; write it simply as B. We have shown that

$$Z = \mathrm{id}_{\mathsf{C}} \otimes B.$$

Since $Z \in \mathfrak{sl}(\mathsf{C} \otimes \mathsf{R})$, we have $B \in \mathfrak{sl}(\mathsf{R})$. Thus

$$\mathrm{Lie}(P) = \mathrm{id}_{\mathsf{C}} \otimes \mathfrak{sl}(\mathsf{R}).$$

Since H_2 is connected and is contained in P, it follows that

$$P^\circ = H_2.$$

Conjugating by g proves (3.4). In particular, $g^{-1} H_2 g$ is defined over $\mathbb{Q}$. Let $\mathcal{P}_d$ be the real vector space of homogeneous degree-d polynomials on $\mathbb{R}^N$. Then the subspace $\mathcal{P}_d^{g^{-1} H_2 g}$ is a $\mathbb{Q}$-defined linear subspace of $\mathcal{P}_d$. We now identify this invariant subspace. Put

$$F_g := \det \circ g.$$

After composing polynomials with g^{-1}, the computation reduces to the standard right $\mathrm{SL}(\mathsf{R})$-action on $\mathsf{C} \otimes \mathsf{R}$. By the Cauchy decomposition,

$$\mathrm{Sym}^d((\mathsf{C} \otimes \mathsf{R})^*) \simeq \bigoplus_{\lambda \vdash d} \mathsf{S}_\lambda(\mathsf{C}^*) \otimes \mathsf{S}_\lambda(\mathsf{R}^*).$$

The restriction of $\mathsf{S}_\lambda(\mathsf{R}^*)$ to $\mathrm{SL}(\mathsf{R})$ contains a trivial subrepresentation if and only if $\lambda = (1^d)$. Hence the space of right-$\mathrm{SL}(\mathsf{R})$-invariant homogeneous degree-d polynomials is

$$\mathsf{S}_{(1^d)}(\mathsf{C}^*) \otimes \mathsf{S}_{(1^d)}(\mathsf{R}^*) = \wedge^d \mathsf{C}^* \otimes \wedge^d \mathsf{R}^*,$$

which is one-dimensional and is spanned by the determinant. Therefore

$$\mathcal{P}_d^{g^{-1} H_2 g} = \mathbb{R} F_g.$$

Since this one-dimensional subspace is defined over $\mathbb{Q}$, the line $\mathbb{R} F_g$ is defined over $\mathbb{Q}$. Equivalently, F_g is proportional to a polynomial with rational coefficients. By Theorem 2.6, the lattice $\Delta = g\mathbb{Z}^N$ is determinant-rational.

The proof for $\mathcal{M}_{k,2}$ is identical, with the left $\mathrm{SL}_d(\mathbb{R})$-factor in place of the right factor. □

Remark 3.4. Indeed, the above proof shows that if $\mathcal{M}_{k,1}$ is rational with respect to $\wedge^{kd} \Delta$, then the individual right factor $g^{-1} H_2 g$ is defined over $\mathbb{Q}$. Thus the corresponding determinant form is of inner type.

Remark 3.5 (The case $d = 2$)*.* When $d = 2$, there is only one critical degree, namely $kd = 2$. The column exceptional summand $\mathcal{M}_{1,1}$ is three-dimensional, and the Plücker lines of column-isotropic 2-planes form a nondegenerate conic in $\mathbb{P}(\mathcal{M}_{1,1})$. If a non-determinant-rational lattice Δ had three distinct Δ-rational column-isotropic planes, their Plücker lines would give three distinct rational points on this conic. Since no line contains three distinct points of a nondegenerate conic, these three Plücker lines would span $\mathbb{P}(\mathcal{M}_{1,1})$. Equivalently, $\mathcal{M}_{1,1}$ would be rational with respect to the exterior lattice $\wedge^2\Delta$, and the proposition would force Δ to be determinant-rational.

Thus a non-determinant-rational lattice in $\mathrm{M}_2(\mathbb{R})$ has at most two rational column-isotropic planes. The same argument gives at most two rational row-isotropic planes. Hence there are at most four maximal rational linear subspaces contained in $\{\det = 0\}$. This is the finiteness phenomenon used in the quadratic determinant case [14].

3.6. **Diophantine lattices and determinant forms.** The exceptional subspaces $\mathcal{M}_{k,1}$ and $\mathcal{M}_{k,2}$ are precisely the directions in which the local contraction estimates may fail. The following condition requires rational Plücker configurations to remain polynomially separated from these subspaces unless they lie in them exactly.

Definition 3.6 (Diophantine lattices)**.** Let $\eta > 0$ and $M > 1$. A lattice $\Lambda < \mathrm{M}_d(\mathbb{R})$ is (η, M)-*Diophantine* if for every $1 \le k \le d-1$, $m \in \{1, 2\}$, $r \in \{1, \dim \mathcal{M}_{k,m}\}$, and for all kd-dimensional Λ-rational subspaces $V_1, \ldots, V_r$, the vector

$$\mathbf{w} := \mathsf{w}_{\Lambda,V_1} \wedge \cdots \wedge \mathsf{w}_{\Lambda,V_r}.$$

satisfies that

$$\text{either } \mathbf{w} = \pi^{(r)}_{k,m}(\mathbf{w}) \quad \text{ or } \quad \|\mathbf{w} - \pi^{(r)}_{k,m}(\mathbf{w})\| \ge \eta \, \|\mathbf{w}\|^{-M}.$$

A lattice is *Diophantine* if it is (η, M)-Diophantine for some $\eta > 0$ and $M > 1$.

The Diophantine property is invariant under left and right multiplication by elements of $\mathrm{GL}_d(\mathbb{R})$, and under transposition. Indeed, these linear transformations preserve the two exceptional families (with transposition interchanging them), and the relevant norms are comparable under every fixed invertible linear map.

Proposition 3.7. *Every lattice in* $\mathrm{M}_d(\mathbb{R})$ *consisting of matrices with algebraic entries is Diophantine.*

The proof is given in Section 13.

Our first counting theorem for general lattices is the following extension of Theorem 1.1.

Theorem 3.8. *Let* $\Lambda < \mathrm{M}_d(\mathbb{R})$ *be a Diophantine lattice that is not determinant-rational. Then, for every* $a < b$*,*

$$N^{\times}_{\Lambda}(a, b; T) \sim \frac{C_{\|\cdot\|}}{\mathrm{covol}(\Lambda)}(b-a)T^{d(d-1)} \qquad (T \to \infty),$$

where $C_{\|\cdot\|}$ is the constant appearing in (1.3)*. Equivalently,*

$$N_\Lambda^\times(a,b;T) \sim \frac{1}{\operatorname{covol}(\Lambda)} \operatorname{vol}\{v \in \mathrm{M}_d(\mathbb{R}) : \ \|v\| < T, \ a < \det v < b\}.$$

We also record the corresponding terminology for determinant forms. If $F = F_{\Lambda,\mathcal{B}}$, where $\mathcal{B}$ is a $\mathbb{Z}$-basis of Λ, let

$$F^* := F_{\Lambda^*,\mathcal{B}^*},$$

where $\mathcal{B}^*$ is the basis dual to $\mathcal{B}$ with respect to the trace pairing.

Definition 3.9. A determinant form F is *Diophantine* if, for one (equivalently, every) realization $F = F_{\Lambda,\mathcal{B}}$, the lattice Λ is Diophantine in the sense of Definition 3.6.

This definition is independent of the realization. Indeed, if $F = \det \circ g = \det \circ g'$, then $g'g^{-1}$ preserves the determinant; by Lemma 2.2, it is given by left and right multiplication, possibly followed by transposition. The preceding invariance observation therefore applies. A change of integral coordinates merely changes the chosen basis of the same lattice. Theorem 3.8 is therefore equivalent to the following:

Theorem 3.10. *Let F be a Diophantine determinant form on* $\mathrm{M}_d(\mathbb{R})$*. If F is not proportional to a polynomial with rational coefficients, then the asymptotic formula of Theorem 1.4 holds for F.*

3.7. **Quasi-null subspaces and the modified height function.** Motivated by Lemma 3.2, we now introduce subspaces whose Plücker vectors are close to one of the exceptional summands.

For $0 \le \eta < 1$, $M \ge 1$, and $1 \le k \le d-1$, let

$$\mathscr{Q}_{kd,\eta,M} := \left\{ 0 \neq v \in \wedge^{kd} \mathrm{M}_d(\mathbb{R}) : \min_{m=1,2} \|v - \pi_{k,m}(v)\| \le \eta \|v\|^{-M} \right\}. \quad (3.5)$$

Definition 3.11. Let $0 \le \eta < 1$, $M \ge 1$, and let $\Delta < \mathrm{M}_d(\mathbb{R})$ be a lattice. For $1 \le k \le d-1$, a kd-dimensional Δ-rational subspace V is (η, M)*-quasi-null* if $\mathsf{w}_{\Delta,V} \in \mathscr{Q}_{kd,\eta,M}$, i.e.

$$\min_{m=1,2} \|\mathsf{w}_{\Delta,V} - \pi_{k,m}(\mathsf{w}_{\Delta,V})\| \le \eta \|\mathsf{w}_{\Delta,V}\|^{-M}.$$

It is *column quasi-null* if the inequality holds with $m = 1$, and *row quasi-null* if it holds with $m = 2$. A subspace may satisfy both conditions.

For a lattice $\Delta \in X$ and $1 \le i \le N-1$, let $\widetilde{\mathscr{Q}}_{i,\eta,M}(\Delta)$ be the collection of all i-dimensional Δ-rational subspaces V for which there exist $1 \le k \le d-1$ and a kd-dimensional Δ-rational subspace V' such that

$$V \subset V' \qquad \text{and} \qquad \mathsf{w}_{\Delta,V'} \in \mathscr{Q}_{kd,\eta,M}.$$

Thus $\widetilde{\mathscr{Q}}_{i,\eta,M}(\Delta)$ is the downward closure, in all dimensions, of the critical quasi-null subspaces. In particular, if $V \notin \widetilde{\mathscr{Q}}_{i,\eta,M}(\Delta)$ and $V \subset W$, where W

is Δ-rational and $1 \le \dim W \le N-1$, then

$$W \notin \widetilde{\mathscr{Q}}_{\dim W,\eta,M}(\Delta).$$

Indeed, if W were contained in a critical quasi-null subspace, then so would V.

Definition 3.12 (Modified Margulis height function)**.** For $h \in H$ and $\Delta \in X$, define

$$\widehat{\alpha}_{\eta,M}(h;\Delta) := \max_{1\le i\le N-1} \widehat{\alpha}_{i,\eta,M}(h;\Delta), \tag{3.6}$$

where

$$\widehat{\alpha}_{i,\eta,M}(h;\Delta) := \max\left\{1, \sup_{\substack{V \text{ is } \Delta\text{-rational, } \dim V=i \\ V\notin\widetilde{\mathscr{Q}}_{i,\eta,M}(\Delta)}} \frac{1}{d_{h\Delta}(hV)}\right\}. \tag{3.7}$$

Here the supremum over the empty set is understood to be 0.

It follows directly from the defining quasi-null inequality that for any $1 \le i \le N-1$ and $0 \le \eta' \le \eta < 1$,

$$\widetilde{\mathscr{Q}}_{i,\eta',M}(\Delta) \subset \widetilde{\mathscr{Q}}_{i,\eta,M}(\Delta), \qquad \widehat{\alpha}_{\eta,M}(h;\Delta) \le \widehat{\alpha}_{\eta',M}(h;\Delta). \tag{3.8}$$

3.8. **Main technical results.** For $t \ge 0$, define

$$b_t := \operatorname{diag}(e^{-t},\dots,e^{-t},e^{(d-1)t}) \in \mathrm{SL}_d(\mathbb{R}), \tag{3.9}$$

and

$$a_t := (b_t, b_t) \in H. \tag{3.10}$$

Let $K = \mathrm{SO}(d) \times \mathrm{SO}(d)$, and let dk denote its probability Haar measure.

The first result is the uniform integrability estimate for the modified height.

Theorem 3.13. *Let $\Lambda \in X$ be an (η, M)-Diophantine lattice, where $0 < \eta < 1$ and $M > 1$. Then there exists $0 < \theta \le (2d)^{-6}$ such that*

$$\sup_{t\ge0} \int_K \widehat{\alpha}_{\eta,M}(a_t k;\Lambda)^{1+\theta}\, dk < \infty.$$

Let Λ_{iso} denote the set of lattice points that lie in a Λ-rational column- or row-isotropic subspace of one of the dimensions $d, 2d, \dots, d(d-1)$.

The singular asymptotic requires the following mild noncoincidence condition.

Definition 3.14. Let $\Delta < \mathrm{M}_d(\mathbb{R})$ be a lattice. For every Δ-rational subspace $V < \mathrm{M}_d(\mathbb{R})$, define the quotient lattice

$$\Delta_V := p_{V^\perp}(\Delta) < V^\perp,$$

where $p_{V^\perp}$ denotes orthogonal projection. This is the Euclidean realization of the abstract quotient $\Delta/(\Delta \cap V)$.

We say that Δ satisfies the *column isotropic noncoincidence condition* if, for every $\lceil d/2 \rceil \le k \le d-2$, there are no distinct subspaces V_1, V_2

that are maximal, with respect to inclusion, among the proper Δ-rational column-isotropic subspaces, have dimension kd, and satisfy

$$h\Delta_{V_1} \quad \text{is commensurable with} \quad \Delta_{V_2}$$

for some $h \in \mathrm{GL}_d(\mathbb{R})$, where h acts by left multiplication on $\mathrm{M}_d(\mathbb{R})$.

We say that Δ satisfies the *row isotropic noncoincidence condition* if the analogous condition holds for row-isotropic subspaces, with

$$\Delta_{V_1}h \quad \text{commensurable with} \quad \Delta_{V_2}$$

for some $h \in \mathrm{GL}_d(\mathbb{R})$, where h acts by right multiplication. Finally, Δ satisfies the *isotropic noncoincidence condition* if it satisfies both the column and row isotropic noncoincidence conditions.

For $d = 3$, we have $\lceil d/2 \rceil = 2$ and $d - 2 = 1$ and hence the above noncoincidence condition is vacuous. See Example **??** for more examples.

For every Δ-rational subspace V, one has

$$\Delta_V^* = V^\perp \cap \Delta^*.$$

Taking dual lattices, the column commensurability condition is equivalent to

$$(h^{-1})^\intercal (V_1^\perp \cap \Delta^*) \quad \text{being commensurable with} \quad V_2^\perp \cap \Delta^*.$$

Since $h \mapsto (h^{-1})^\intercal$ is a bijection of $\mathrm{GL}_d(\mathbb{R})$, it is equivalently the nonexistence of $g \in \mathrm{GL}_d(\mathbb{R})$ such that

$$g(V_1^\perp \cap \Delta^*) \quad \text{is commensurable with} \quad V_2^\perp \cap \Delta^*.$$

Likewise, the row condition is equivalently the nonexistence of $g \in \mathrm{GL}_d(\mathbb{R})$ such that

$$(V_1^\perp \cap \Delta^*)g \quad \text{is commensurable with} \quad V_2^\perp \cap \Delta^*.$$

Theorem 3.15. *Let $\Lambda < \mathrm{M}_d(\mathbb{R})$ be a Diophantine lattice that satisfies the isotropic noncoincidence condition. Then the limit*

$$c_\Lambda^{\mathrm{sing}} := \lim_{T\to\infty} T^{-d(d-1)} \#\{v \in \Lambda_{\mathrm{iso}} : \|v\| \le T\}$$

exists and is finite. In particular, $c_\Lambda^{\mathrm{sing}} \ge 0$.

4. Sublevel estimates

The goal of this section is to prove the uniform sublevel estimate needed for the local contraction inequalities. We work one Schur module at a time. Along the principal diagonal ray, the module decomposes into b_t-weight spaces, and the expanding horospherical subgroup acts triangularly with respect to this decomposition. Thus the coordinate at the first nonnegative weight is a vector-valued polynomial in the horospherical parameter $\xi \in \mathbb{R}^{d-1}$.

For $d \ge 3$, this polynomial may have degree greater than one. We prove the required sublevel estimate by combining codimension bounds for its homogeneous initial forms, obtained from the semistandard Young tableau basis, with analytic stability results of Cluckers–Miller. The equality cases in the codimension bounds are precisely the rectangular representations, which later give the exceptional isotropic summands.

Let

$$b_t := \operatorname{diag}(e^{-t}, \dots, e^{-t}, e^{(d-1)t}) \in \mathrm{SL}_d(\mathbb{R}).$$

Let $U < \mathrm{SL}_d(\mathbb{R})$ be the expanding horospherical subgroup for b_t:

$$U = \{g \in \mathrm{SL}_d(\mathbb{R}) : b_{-t} g b_t \to I_d \text{ as } t \to +\infty\}.$$

Equivalently,

$$U = \left\{ u_\xi := \begin{pmatrix} \mathrm{Id}_{d-1} & 0 \\ \xi & 1 \end{pmatrix} : \xi \in \mathbb{R}^{d-1} \right\}.$$

We identify its Lie algebra with

$$\mathfrak{u} \simeq \sum_{1 \le j \le d-1} \mathbb{R} E_{dj} \simeq \mathbb{R}^{d-1},$$

and for $\xi \in \mathbb{R}^{d-1}$ we write

$$Y_\xi := \log u_\xi = \sum_{j=1}^{d-1} \xi_j E_{dj} \in \mathfrak{u}.$$

Thus $u_\xi = \exp(Y_\xi)$, and the map $\xi \mapsto Y_\xi$ is linear.

In this section we fix

$$V = \mathsf{S}_\lambda(\mathbb{R}^d),$$

where λ is a partition whose Young diagram fits inside the $d \times d$ box. Since we work with $\mathrm{SL}_d(\mathbb{R})$-representations, we may replace λ by the equivalent partition $\lambda - \lambda_d(1^d)$, and we shall therefore assume throughout that

$$\lambda_d = 0. \tag{4.1}$$

In particular, the Young diagram of λ has at most $d-1$ nonzero rows and at most d columns.

Put

$$\mu_i := -|\lambda| + di \quad (0 \le i \le \lambda_1),$$

and let V_i be the $\log b_1$-eigenspace of weight μ_i. A nonzero vector contained in some V_i will be called a weight homogeneous vector. We have

$$V = \bigoplus_{i=0}^{\lambda_1} V_i.$$

Indeed, in the semistandard-tableau model, the index i is the number of entries equal to d. Moreover,

$$[\log b_1, E_{dj}] = dE_{dj}, \qquad \text{so} \qquad Y_\xi V_i \subset V_{i+1}.$$

Let

$$L = \begin{pmatrix} \mathrm{SL}_{d-1}(\mathbb{R}) & 0 \\ 0 & 1 \end{pmatrix}.$$

Since L commutes with b_t, every V_i is L-invariant. Let

$$p_i : V \to V_i$$

denote the projection onto V_i. Because $\mathfrak{u}$ is abelian, for every $v \in V$,

$$u_\xi v = \sum_{k \geq 0} \frac{1}{k!} Y_\xi^k v.$$

Since Y_ξ^k maps V_i to V_{i+k}, we obtain, for $0 \leq m \leq \lambda_1$,

$$p_m(u_\xi v) = \sum_{i=0}^{m} \frac{1}{(m-i)!} Y_\xi^{m-i} p_i(v). \tag{4.2}$$

In particular, $p_m(u_\xi v)$ is a polynomial in ξ of degree at most m.

The goal of this section is the following uniform sublevel estimate.

Proposition 4.1 (Sublevel estimate)**.** *Let $1 \leq \ell = \ell(\lambda) \leq \lambda_1$ be the smallest integer such that $\mu_\ell \geq 0$. Let $\Omega \subset \mathbb{R}^{d-1}$ be a bounded open set, and define*

$$A_\ell(v) := \max_{0 \leq i \leq \ell} \|p_i(v)\|.$$

Put

$$\alpha_\lambda := \begin{cases} \dfrac{d-1}{d-2}, & \text{if } \mu_\ell = 0 \text{ and } \lambda \neq (d^k),\ 1 \leq k \leq d-1, \\ 1, & \text{otherwise.} \end{cases}$$

Then for every $\eta > 0$ there exists $C = C(\eta, \Omega) > 0$ such that, for every $v \in V$ with $A_\ell(v) > 0$ and every $0 < \varepsilon < 1$,

$$\mathrm{Leb}\Big(\big\{\xi \in \Omega : \|p_\ell(u_\xi v)\| \leq \varepsilon A_\ell(v)\big\}\Big) \leq C \varepsilon^{\alpha_\lambda - \eta}. \tag{4.3}$$

Remark 4.2 (The case $d = 2$)*.* For $d = 2$, after reducing by determinant twists, the first case in the definition of α_λ never arises: for every nonexceptional λ, one has $\ell = 1$ and $\mu_\ell > 0$, while the case $\mu_\ell = 0$ is precisely the rectangular exceptional case $\lambda = (2) = (d^1)$. Thus the analogue of the proposition gives the exponent $1 - \eta$. In fact, one obtains the sharper exponent 1, because $p_1(u_\xi v) = p_1(v) + Y_\xi p_0(v)$ is affine-linear in the one horospherical parameter.

4.1. **Tableaux and a codimension bound.** We use the standard semistandard Young tableau basis of the Schur module $\mathsf{S}_\lambda(\mathbb{R}^d)$; see [15, Chapters 1 and 8]. A semistandard Young tableau of shape λ with entries in $\{1, \ldots, d\}$ is a filling of the Young diagram of λ by numbers in $\{1, \ldots, d\}$, weakly increasing along rows and strictly increasing down columns. The Schur module $\mathsf{S}_\lambda(\mathbb{R}^d)$ has a basis indexed by such tableaux. We also use the standard-basis theorem and the triangular straightening law for Schur modules; see [1] and [19, Chapter 2].

If T is a semistandard Young tableau, write

$$n_j(\mathsf{T}) := \#\{\text{entries of } \mathsf{T} \text{ equal to } j\}.$$

Its content is the d-tuple

$$(n_1(\mathsf{T}), \ldots, n_d(\mathsf{T})).$$

Since

$$\log b_1 = \mathrm{diag}(-1, \ldots, -1, d-1),$$

the $\log b_1$-weight of T is

$$\mathrm{wt}(\mathsf{T}) = -\sum_{j=1}^{d-1} n_j(\mathsf{T}) + (d-1)n_d(\mathsf{T}) = -|\lambda| + dn_d(\mathsf{T}). \tag{4.4}$$

For $1 \le i \le d-1$, write

$$Y_i := Y_{e_i} = E_{di}.$$

Thus Y_i sends e_i to e_d and kills e_j for $j \neq i$. In particular, Y_i decreases the i-content by one, increases the d-content by one, and raises the $\log b_1$-weight by d.

We say that an entry $i < d$ of a tableau T is i-free if the column containing that entry contains no d. Equivalently, replacing that entry by d does not create two equal entries in the same column.

Fix once and for all the standard semistandard Young tableau basis of $\mathsf{S}_\lambda(\mathbb{R}^d)$, together with a standard monomial order on this basis. If

$$x = \sum_{\mathsf{T}} c_{\mathsf{T}} \mathsf{T},$$

where T ranges over this basis, we write

$$\mathrm{supp}(x) := \{\mathsf{T} : c_{\mathsf{T}} \neq 0\}.$$

Lemma 4.3 (Triangularity of the raising operators)**.** *Let $0 \neq x \in \mathsf{S}_\lambda(\mathbb{R}^d)$, and write*

$$x = \sum_{\mathsf{T}} c_{\mathsf{T}} \mathsf{T}$$

in the semistandard Young tableau basis. Let T_0 be the largest tableau in $\mathrm{supp}(x)$. *Then, for every $1 \le i \le d-1$,*

$$Y_i x = 0 \quad \Longrightarrow \quad Y_i \mathsf{T}_0 = 0.$$

Equivalently, if $Y_i x = 0$, then T_0 has no i-free entry.

Proof. The operator $Y_i = E_{di}$ acts by replacing one occurrence of i by d, followed by straightening. Such a replacement is nonzero precisely when the chosen occurrence is i-free. The standard-basis theorem and the triangular straightening law imply that, whenever $Y_i\mathsf{T}_0 \neq 0$, the largest standard tableau occurring in $Y_i\mathsf{T}_0$ is strictly larger than every tableau that can arise from $Y_i\mathsf{T}$ with $\mathsf{T} < \mathsf{T}_0$; see [1] and [19, Chapter 2]. Its coefficient therefore cannot cancel in $Y_i x$. Thus $Y_i x = 0$ forces $Y_i \mathsf{T}_0 = 0$, as claimed. ☐

For a $\log b_1$-weight homogeneous vector v, and for $r \ge 1$, set

$$\mathcal{Z}_r(v) := \{\xi \in \mathbb{R}^{d-1} : Y_\xi^r v = 0\}. \tag{4.5}$$

Since Y_ξ depends linearly on ξ, the map $\xi \mapsto Y_\xi^r v$ has polynomial coordinates. Hence $\mathcal{Z}_r(v)$ is a real algebraic, in particular semialgebraic, subset of

$\mathbb{R}^{d-1}$. We write $\dim \mathcal{Z}_r(v)$ for its semialgebraic dimension, equivalently the maximum dimension of a smooth semialgebraic stratum, and define

$$\operatorname{codim} \mathcal{Z}_r(v) := (d-1) - \dim \mathcal{Z}_r(v).$$

Also define

$$m(v) := \min\{m \geq 0 : \mathrm{wt}(v) + dm \geq 0\}. \tag{4.6}$$

Equivalently, if $\mathrm{wt}(v) < 0$, then

$$m(v) = \left\lceil \frac{-\,\mathrm{wt}(v)}{d} \right\rceil.$$

Lemma 4.4 (Codimension bound)**.** *Let $v \in \mathsf{S}_\lambda(\mathbb{R}^d)$ be $\log b_1$-weight homogeneous. For $1 \leq r \leq m(v)$,*

$$\operatorname{codim} \mathcal{Z}_r(v) \geq m(v) + 1 - r.$$

Moreover, if $d \mid |\lambda|$ and equality holds for some $1 \leq r \leq m(v)$, then

$$\lambda = (d^k) \qquad \text{for some } 1 \leq k \leq d-1.$$

Proof. We argue by induction on r. Choose a smooth semialgebraic stratum X of $\mathcal{Z}_r(v)$ of maximal dimension, and write

$$\mathsf{s} := \dim X = \dim \mathcal{Z}_r(v), \qquad \mathsf{c} := (d-1) - \mathsf{s} = \operatorname{codim} \mathcal{Z}_r(v).$$

If $r > 1$ and $X \subset \mathcal{Z}_{r-1}(v)$, then the induction hypothesis gives

$$\mathsf{c} \geq \operatorname{codim} \mathcal{Z}_{r-1}(v) \geq m(v) + 2 - r,$$

which is stronger than required. Hence we may assume that either $r = 1$ or $X \not\subset \mathcal{Z}_{r-1}(v)$. Now choose $\xi \in X$, which we also assume $\xi \notin \mathcal{Z}_{r-1}(v)$ if $r > 1$. Then

$$Y_\xi^r v = 0 \quad \text{and} \quad Y_\xi^{r-1} v \neq 0.$$

Let $E := T_\xi X$; so $\dim E = s$. For $\eta \in E$, choose a C^1-curve $\xi_\eta(t)$ in X with $\xi_\eta(0) = \xi$ and $\xi_\eta'(0) = \eta$. Differentiating $Y_{\xi_\eta(t)}^r v = 0$ at $t = 0$, and using the commutativity of $\mathfrak{u}$, gives

$$0 = r Y_\eta Y_\xi^{r-1} v.$$

Set $w := Y_\xi^{r-1} v$. Then $w \neq 0$, the vector w is weight homogeneous, and $Y_\eta w = 0$ for any $\eta \in E$.

Choose $h_0 \in \mathrm{SL}_{d-1}(\mathbb{R})$ such that $E h_0^{-1} = \langle e_1, \ldots, e_s \rangle$, and put $h = \operatorname{diag}(h_0, 1) \in L$. With our row-vector convention for ξ, direct matrix multiplication gives $h Y_\eta h^{-1} = Y_{\eta h_0^{-1}}$. Since h commutes with b_1, the vector hw has the same weight as w. Replacing w by hw, we may therefore assume that

$$Y_i w = 0 \qquad (1 \leq i \leq \mathsf{s}). \tag{4.7}$$

Let T be the largest semistandard Young tableau occurring in w. By Lemma 4.3, (4.7) implies that T has no i-free entry for $1 \leq i \leq \mathsf{s}$. Thus every free entry of T has label in the set $\{\mathsf{s}+1, \ldots, d-1\}$ which has cardinality $\mathsf{c} = d - 1 - \mathsf{s}$. The Young diagram has at most d columns, and a fixed label occurs at most once in each column. Hence, if $N_{\text{free}}(\mathsf{T})$ denotes the number of free entries, then

$$N_{\text{free}}(\mathsf{T}) \leq d\mathsf{c}. \tag{4.8}$$

Let $N_{\text{nonfree}}(\mathsf{T})$ be the number of entries less than d in columns containing a d. Because $\lambda_d = 0$, every column has height at most $d-1$. A column containing a d therefore contains at most $d-2$ other entries. Since a column contains at most one d,

$$N_{\text{nonfree}}(\mathsf{T}) \leq (d-2)n_d(\mathsf{T}).$$

Using (4.4), we obtain

$$\begin{aligned} \text{wt}(\mathsf{T}) &= -N_{\text{free}}(\mathsf{T}) - N_{\text{nonfree}}(\mathsf{T}) + (d-1)n_d(\mathsf{T}) \\ &\geq -N_{\text{free}}(\mathsf{T}) + n_d(\mathsf{T}) \geq -d\mathsf{c}. \end{aligned}$$

Since T occurs in the weight-homogeneous vector w,

$$\text{wt}(w) = \text{wt}(\mathsf{T}) \geq -d\mathsf{c}. \tag{4.9}$$

On the other hand, $\text{wt}(w) = \text{wt}(v) + d(r-1)$. The definition of $m(v)$ gives $0 \leq \text{wt}(v) + dm(v) < d$, and hence

$$0 \leq \text{wt}(w) + d(m(v) - r + 1) < d. \tag{4.10}$$

Combining (4.9) and (4.10), we find

$$-d\mathsf{c} \leq \text{wt}(w) < -d(m(v) - r).$$

Thus $\mathsf{c} > m(v) - r$, and, since c is an integer,

$$\mathsf{c} \geq m(v) + 1 - r.$$

This proves the codimension estimate.

Assume now that $d \mid |\lambda|$ and that equality holds. Then $\mathsf{c} = m(v) + 1 - r$. Every weight of $\mathsf{S}_\lambda(\mathbb{R}^d)$ is congruent to $-|\lambda|$ modulo d, so $\text{wt}(w) \in d\mathbb{Z}$. The inequalities above become

$$-dc \leq \text{wt}(w) < -d(\mathsf{c} - 1),$$

and therefore $\text{wt}(w) = -d\mathsf{c}$.

The sharper tableau estimate

$$\text{wt}(\mathsf{T}) \geq -N_{\text{free}}(\mathsf{T}) + n_d(\mathsf{T}) \geq -d\mathsf{c} + n_d(\mathsf{T})$$

then forces

$$n_d(\mathsf{T}) = 0, \qquad N_{\text{free}}(\mathsf{T}) = d\mathsf{c}.$$

Thus every entry of T is free and has label in $\{\mathsf{s}+1, \ldots, d-1\}$. There are c possible labels, each occurring at most once in each of at most d columns. Equality in (4.8) consequently forces the diagram to have exactly d columns, each containing all c labels. Hence $\lambda = (d^{\mathsf{c}})$, which proves the equality statement. $\square$

4.2. **Uniform negative moments and initial forms.** We record two analytic facts for compact semialgebraic families of polynomial maps. The external input is the one-parameter stability statement Lemma A.1, a consequence of the preparation and constructible-integration theory of Cluckers–Miller in the globally subanalytic category. Since semialgebraic sets and polynomial maps are globally subanalytic, this input applies to the families considered below.

The first lemma upgrades pointwise integrability of the fibers to a uniform negative-moment bound, after a small loss in the exponent. The second uses this uniformity to pass from integrability of homogeneous initial forms at zeros to integrability of the full polynomial family.

Lemma 4.5 (Uniform negative moments in compact polynomial families)**.** *Let $A \subset \mathbb{R}^N$ be a compact semialgebraic set, let $B \subset \mathbb{R}^n$ be a bounded semialgebraic open set, and let W be a finite-dimensional Euclidean space. Let*

$$Q : A \times B \to W, \qquad Q_a(x) := Q(a, x),$$

be the restriction of a polynomial map, and put $F(a, x) := \|Q_a(x)\|^2$. Assume that, for some $s_1 > 0$,

$$F(a, \cdot)^{-s_1} \in L^1(B) \qquad \text{for every } a \in A. \tag{4.11}$$

Then, for every $0 < s < s_1$,

$$\sup_{a \in A} \int_B F(a, x)^{-s}\, dx < \infty. \tag{4.12}$$

Consequently,

$$\mathrm{Leb}\{x \in B : \|Q_a(x)\| \le \varepsilon\} \ll_{A,B,Q,s} \varepsilon^{2s} \qquad (0 < \varepsilon < 1),$$

uniformly in $a \in A$.

Proof. Fix $0 < s < s_1$. Choose a rational number σ and a real number τ such that

$$s < \sigma < \tau < s_1.$$

Since F is bounded above on $A \times B$, there is a constant C_0 such that

$$F(a, x)^{-s} \le C_0 F(a, x)^{-\sigma} \qquad (F(a, x) > 0). \tag{4.13}$$

Moreover, $F(a, x)^{-\tau} \le 1 + F(a, x)^{-s_1}$. Thus

$$F(a, \cdot)^{-\tau} \in L^1(B) \qquad (a \in A), \tag{4.14}$$

and it is enough to prove that the σ-moments are uniformly bounded.

Put

$$\mathcal{D} := \{(a, x) \in A \times B : F(a, x) > 0\},$$

and extend $F^{-\sigma}|_{\mathcal{D}}$ by zero to $A \times B$. Since $\sigma \in \mathbb{Q}$, this extension is globally subanalytic. By the stability of constructible functions under parameterized integration [8, Theorem 1.3],

$$I_\sigma(a) := \int_B F(a, x)^{-\sigma}\, dx \tag{4.15}$$

is a constructible function on A. The value assigned to the integrand on $F^{-1}(0)$ is immaterial: by the hypothesis, no polynomial $F(a,\cdot)$ is identically zero, and hence its zero set has Lebesgue measure zero.

We claim that I_σ is locally bounded on A. Suppose not. Then the preparation theorem for constructible functions, together with the no-cancellation statement of Cluckers–Miller [8, Theorem 3.11 and Proposition 6.1], gives a globally subanalytic arc

$$a : [0,\delta) \to A, \qquad a(0) = a_0,$$

which, after a power reparameterization, is real analytic at 0, and such that

$$I_\sigma(a(t)) \to \infty \qquad (t \to 0). \tag{4.16}$$

Apply the one-parameter stability statement to

$$H(t,x) := F(a(t),x).$$

The function H is real analytic and globally subanalytic near $[0,\delta) \times \overline{B}$, and (4.14) gives

$$H(0,\cdot)^{-\tau} \in L^1(B).$$

Therefore by Lemma A.1, there exists $\delta' > 0$ such that

$$\sup_{0\le t<\delta'} I_\sigma(a(t)) < \infty.$$

This contradicts (4.16). Hence I_σ is locally bounded. Since A is compact,

$$\sup_{a\in A} I_\sigma(a) < \infty.$$

Together with (4.13), this proves

$$\sup_{a\in A} \int_B F(a,x)^{-s}\, dx < \infty.$$

Finally, Chebyshev's inequality gives

$$\mathrm{Leb}\{x \in B : \|Q_a(x)\| \le \varepsilon\} \le \varepsilon^{2s} \int_B \|Q_a(x)\|^{-2s}\, dx,$$

and the asserted sublevel estimate follows. □

Lemma 4.6 (Initial forms at all nearby zeros control local integrability)**.** *Let $B \subset \mathbb{R}^n$ be a bounded semialgebraic open set, let W be a finite-dimensional Euclidean space, and let $\mathcal{C}$ be a compact semialgebraic family of nonzero polynomial maps $Q : \mathbb{R}^n \to W$ of degree at most D. Suppose that there exists $\alpha > 0$ such that, for every $Q \in \mathcal{C}$ and every zero $x_0 \in \overline{B}$ of Q, the first nonzero homogeneous Taylor term $H_{Q,x_0} := \mathrm{in}_{x_0} Q$ satisfies*

$$\int_{\|y\|<1} \|H_{Q,x_0}(y)\|^{-\gamma}\, dy < \infty \qquad \text{for every } 0 < \gamma < \alpha.$$

Then, for every $0 < \beta < \alpha$,

$$\sup_{Q\in\mathcal{C}} \int_B \|Q(x)\|^{-\beta}\, dx < \infty.$$

Consequently,

$$\mathrm{Leb}\{x \in B : \|Q(x)\| \le \varepsilon\} \ll_{\mathcal{C},B,\beta} \varepsilon^{\beta}$$

uniformly for $Q \in \mathcal{C}$ *and* $0 < \varepsilon < 1$.

Proof. Fix $0 < \beta < \alpha$. The joint zero set

$$\mathcal{Z} := \{(Q,x) \in \mathcal{C} \times \overline{B} : Q(x) = 0\}$$

is compact and semialgebraic. Choose a finite semialgebraic Whitney stratification of $\mathcal{Z}$ such that the projection to $\mathcal{C}$ is a stratified submersion, and refine it so that the order of vanishing in the x-variables is constant on each source stratum. After a finite refinement of the target stratification, the nonempty intersections of the source strata with the fibers

$$\mathcal{Z}_Q := \{x \in \overline{B} : Q(x) = 0\}$$

form Whitney stratifications satisfying the frontier condition. Hardt triviality gives a uniform bound for the number of fiber strata; see [4, Chapter 9]. Let N be such a bound. Choose exponents

$$\beta = \gamma_0 < \gamma_1 < \cdots < \gamma_{N+1} < \alpha. \tag{4.17}$$

We first fix $Q \in \mathcal{C}$. Order the nonempty strata $S_1, \ldots, S_M$, where $M \le N$, so that a stratum whose closure contains another stratum occurs first. We prove inductively that, near every point of S_j, the function $\|Q\|^{-\delta_j}$ is locally integrable, where $\delta_j := \gamma_{N+1-j}$. For the induction step, put

$$\delta_j^+ := \gamma_{N+2-j} > \delta_j.$$

Every stratum preceding S_j has already been treated at an exponent at least δ_j^+.

Fix $x_0 \in S_j$. After shrinking to a relatively compact coordinate patch in S_j, choose a compact semialgebraic neighborhood $K \subset S_j$ of x_0. The frontier condition allows the ambient neighborhood to be chosen so that every zero of Q in it either lies in S_j or in a stratum preceding S_j. Let $c = n - \dim S_j$. A semialgebraic tubular neighborhood provides a compact semialgebraic set K', a parametrization $\sigma : K' \to K$, a smooth semialgebraic normal frame $\nu_1, \ldots, \nu_c$, and $r_0 > 0$ such that

$$\Phi(u,t) := \sigma(u) + \sum_{i=1}^{c} t_i \nu_i(u)$$

is a diffeomorphism on $K' \times B^c_{r_0}$ onto its image. Put

$$q(u,t) := Q(\Phi(u,t)).$$

The order of normal vanishing is constant on the stratum, say equal to $m \ge 1$, so

$$q(u,t) = A(u,t) + R(u,t), \qquad \|R(u,t)\| \ll \|t\|^{m+1}, \tag{4.18}$$

where $A(u,\cdot)$ is nonzero and homogeneous of degree m.

For $u_0 \in K'$, set $x = \Phi(u_0, 0)$ and $L = D\Phi_{(u_0,0)}$. Comparing the lowest nonzero powers of r in $Q\big(\Phi(u_0 + r\xi, rt)\big)$ gives

$$H_{Q,x}(L(\xi, t)) = A(u_0, t). \tag{4.19}$$

Thus the hypothesis on $H_{Q,x}$, followed by an invertible linear change of variables and Fubini's theorem, shows that

$$\int_{\|t\|<1} \|A(u_0, t)\|^{-\delta_j^+}\, dt < \infty \qquad (u_0 \in K').$$

The coefficient vectors of the maps $A(u, \cdot)$, $u \in K'$, form a compact semialgebraic set. Identifying this set with the parameter space and applying Lemma 4.5 to the polynomial evaluation map yields

$$\sup_{u \in K'} \int_{\|t\|<1} \|A(u, t)\|^{-\delta_j}\, dt < \infty. \tag{4.20}$$

In particular, radial integration for the homogeneous map $A(u_0, \cdot)$ gives

$$c - m\delta_j^+ > 0, \qquad \text{and hence } c - m\delta_j > 0. \tag{4.21}$$

Let

$$\mathcal{A} := \{y \in \mathbb{R}^c : 1/2 < \|y\| < 2\},$$

and define, for $0 \le r \le r_0/2$,

$$q_{u,r}(y) := \begin{cases} r^{-m} q(u, ry), & r > 0, \\ A(u, y), & r = 0. \end{cases}$$

The coefficients of $q_{u,r}$ depend continuously and semialgebraically on (u, r). Their coefficient vectors therefore form a compact semialgebraic set. With this coefficient set as parameter space, evaluation on $\mathcal{A}$ is polynomial. At $r = 0$, the power $-\delta_j^+$ is integrable by the initial-form hypothesis. If $r > 0$, every zero of $q_{u,r}$ corresponds under Φ to a zero of Q outside S_j, hence to a point of a preceding stratum. By the induction hypothesis, $\|Q\|^{-\delta_j^+}$, and therefore $\|q_{u,r}\|^{-\delta_j^+}$, is locally integrable at each such zero. Since the integrand is bounded away from the zero set, a finite cover of $\overline{\mathcal{A}}$ then gives

$$\int_{\mathcal{A}} \|q_{u,r}(y)\|^{-\delta_j^+}\, dy < \infty \qquad (u \in K',\ 0 \le r \le r_0/2).$$

Applying Lemma 4.5 again gives

$$\sup_{u \in K',\, 0 \le r \le r_0/2} \int_{\mathcal{A}} \|q_{u,r}(y)\|^{-\delta_j}\, dy < \infty. \tag{4.22}$$

For $r > 0$, the change of variables $t = ry$ and (4.22) imply

$$\sup_{u \in K'} \int_{r\mathcal{A}} \|q(u, t)\|^{-\delta_j}\, dt \ll r^{c - m\delta_j}.$$

The exponent is positive by (4.21); summing over dyadic annuli proves

$$\sup_{u \in K'} \int_{\|t\|<r_0/2} \|q(u, t)\|^{-\delta_j}\, dt < \infty.$$

The Jacobian of Φ is bounded above and below on a smaller compact tubular neighborhood. Hence $\|Q\|^{-\delta_j}$ is integrable near K, completing the induction over the strata.

Since $M \le N$, every exponent used in the induction is at least γ_1. After all strata have been treated, $\|Q\|^{-\gamma_1}$ is locally integrable near every point of Z_Q. Since Z_Q is compact and $\|Q\|^{-\gamma_1}$ is bounded away from the zero set,

$$\int_B \|Q(x)\|^{-\gamma_1}\, dx < \infty \qquad (Q \in \mathcal{C}). \tag{4.23}$$

Finally, identify $\mathcal{C}$ with its compact semialgebraic coefficient set. The evaluation map is polynomial, so Lemma 4.5, applied with $s_1 = \gamma_1/2$ and $s = \beta/2$, turns (4.23) into

$$\sup_{Q\in\mathcal{C}} \int_B \|Q(x)\|^{-\beta}\, dx < \infty.$$

Chebyshev's inequality gives the final sublevel estimate. □

4.3. **The critical weight-vector sublevel estimate.** We now prove the critical sublevel estimate for $\log b_1$-weight homogeneous vectors.

Lemma 4.7 (Critical sublevel estimate for weight vectors)**.** *Let $v \in \mathsf{S}_\lambda(\mathbb{R}^d)$ be $\log b_1$-weight homogeneous. Assume $m(v) \ge 1$, and define*

$$\alpha(v) := \begin{cases} \dfrac{d-1}{d-2}, & \text{if } \operatorname{wt}(v) + dm(v) = 0 \text{ and } \lambda \ne (d^k),\ 1 \le k \le d-1, \\ 1, & \text{otherwise.} \end{cases}$$

Then for every bounded open set $\Omega \subset \mathbb{R}^{d-1}$ and every $\eta > 0$, there exists $C = C(\eta, \Omega, v) > 0$ such that for every $0 < \varepsilon < 1$,

$$\operatorname{Leb}\{\xi \in \Omega : \|Y_\xi^{m(v)} v\| \le \varepsilon \|v\|\} \le C\varepsilon^{\alpha(v)-\eta}.$$

Proof. By homogeneity in v, we may assume $\|v\| = 1$. For simplicity, set $m = m(v)$. It is enough to prove that, for every $0 < \beta < \alpha(v)$, one has

$$\int_\Omega \|Y_\xi^m v\|^{-\beta}\, d\xi < \infty. \tag{4.24}$$

Indeed, Chebyshev's inequality then gives the desired sublevel estimate after choosing β with

$$\alpha(v) - \eta < \beta < \alpha(v)$$

when $0 < \eta < \alpha(v)$; if $\eta \ge \alpha(v)$, the result is trivial after enlarging the constant.

We prove (4.24) by induction on $m = m(v)$. The case $m = 0$ is immediate. Assume $m \ge 1$, and that the same negative-moment statement has been proved for every weight-homogeneous vector w with $m(w) < m$. By Lemma 4.4, applied with $r = m$, the polynomial map

$$P(\xi) := Y_\xi^m v$$

is not identically zero. Away from the zero set of P, the integrand in (4.24) is bounded. It therefore suffices to prove local integrability of $\|P(\xi)\|^{-\beta}$ near each zero of P.

Let ξ_0 be a zero of P. We divide our proof into two cases.

Case 1: $\xi_0 \notin \mathcal{Z}_1(v)$**.** Choose a bounded semialgebraic open neighborhood B_0 of ξ_0 such that

$$\overline{B_0} \cap \mathcal{Z}_1(v) = \varnothing.$$

We claim that the hypotheses of Lemma 4.6 hold for the singleton family $\{P\}$ on B_0, with exponent $\alpha(v)$.

Let $z \in \overline{B_0}$ be a zero of P. If $m = 1$, then $P(z) = Y_z v = 0$ would imply $z \in \mathcal{Z}_1(v)$, contrary to the choice of B_0. Thus this case is empty when $m = 1$, and we may assume $m \geq 2$. Let j be the largest integer $0 \leq j \leq m-1$ such that $Y_z^j v \neq 0$. Because $z \notin \mathcal{Z}_1(v)$, one has $j \geq 1$. Set

$$q := m - j, \qquad w := Y_z^j v.$$

Then $1 \leq q < m$, and w is $\log b_1$-weight homogeneous. Moreover,

$$\mathrm{wt}(w) + dq = \mathrm{wt}(v) + dm. \tag{4.25}$$

It follows that

$$m(w) = q \tag{4.26}$$

and $\alpha(w) = \alpha(v)$.

Using the commutativity of $\mathfrak{u}$,

$$P(z+\eta) = Y_{z+\eta}^m v = \sum_{k=0}^{m} \binom{m}{k} Y_z^{m-k} Y_\eta^k v. \tag{4.27}$$

By the choice of $j = m - q$, all terms of degree $< q$ in η vanish, and the degree-q term is

$$\binom{m}{q} Y_\eta^q w.$$

This polynomial is not identically zero: by (4.26) and Lemma 4.4, the set $\mathcal{Z}_q(w)$ has positive codimension. Consequently the displayed term is the initial form, namely

$$H_{P,z}(\eta) = \binom{m}{q} Y_\eta^q w. \tag{4.28}$$

By the induction hypothesis applied to w, for every $0 < \beta' < \alpha(w) = \alpha(v)$,

$$\int_{\|\eta\|<1} \|Y_\eta^q w\|^{-\beta'} \, d\eta < \infty.$$

This is exactly the negative-moment induction hypothesis. Thus every zero $z \in \overline{B_0}$ has an initial form with integrability exponent at least $\alpha(v)$. Lemma 4.6 applied to the singleton family $\{P\}$ on B_0 therefore yields

$$\|P(\xi)\|^{-\beta} \in L^1_{\mathrm{loc}}(\xi_0).$$

Case 2: $\xi_0 \in \mathcal{Z}_1(v)$**.** Since $Y_{\xi_0}v = 0$, one has $Y_{\xi_0}^a v = 0$ for all $a \geq 1$. Using the commutativity of $\mathfrak{u}$,

$$P(\xi_0 + \eta) = (Y_{\xi_0} + Y_\eta)^m v = Y_\eta^m v = P(\eta).$$

It is therefore enough to prove local integrability at the origin.

The space $\mathcal{Z}_1(v) = \{\xi \in \mathbb{R}^{d-1} : Y_\xi v = 0\}$ is linear. Let

$$N := \mathcal{Z}_1(v)^\perp, \qquad \mathsf{c} := \dim N = \operatorname{codim} \mathcal{Z}_1(v).$$

Write

$$\xi = x + y, \qquad x \in \mathcal{Z}_1(v), \quad y \in N.$$

Because $Y_x v = 0$, one has

$$P(x + y) = Y_{x+y}^m v = Y_y^m v. \tag{4.29}$$

Thus P factors through the quotient $\mathbb{R}^{d-1}/\mathcal{Z}_1(v)$, and in the orthogonal decomposition $\mathbb{R}^{d-1} = \mathcal{Z}_1(v) \oplus N$ local integrability near 0 reduces to local integrability on the normal space N.

We first show that

$$\int_{S_N} \|Y_\theta^m v\|^{-\beta} \, d\theta < \infty, \tag{4.30}$$

where S_N denotes the unit sphere in N.

Let $\theta_0 \in S_N$ satisfy $Y_{\theta_0}^m v = 0$. Since $\theta_0 \in N = \mathcal{Z}_1(v)^\perp$ and $\theta_0 \neq 0$, we have $\theta_0 \notin \mathcal{Z}_1(v)$. Thus θ_0 falls under Case 1. Choose a bounded semialgebraic open neighborhood B_{θ_0} of θ_0 disjoint from $\mathcal{Z}_1(v)$. By Case 1, $\|P\|^{-\beta}$ is integrable on a smaller neighborhood of θ_0. In the orthogonal decomposition $\mathbb{R}^{d-1} = \mathcal{Z}_1(v) \oplus N$, the function P depends only on the N-coordinate. Fubini's theorem therefore gives local integrability in the N-variable. On an annular neighborhood of θ_0, polar coordinates in N have a smooth Jacobian bounded above and below, so this is equivalent to local integrability of

$$\theta \to \|Y_\theta^m v\|^{-\beta}$$

on S_N near θ_0. The zero set on S_N is compact, while the integrand is bounded away from it. A finite cover proves (4.30).

Now use polar coordinates in N. Since P is homogeneous of degree m,

$$Y_{r\theta}^m v = r^m Y_\theta^m v.$$

For sufficiently small $\delta > 0$,

$$\int_{\|y\|<\delta,\ y \in N} \|Y_y^m v\|^{-\beta} \, dy = \int_0^\delta r^{\mathsf{c}-1-m\beta} \, dr \int_{S_N} \|Y_\theta^m v\|^{-\beta} \, d\theta. \tag{4.31}$$

The angular integral is finite by (4.30). Thus the normal integral is finite provided

$$\mathsf{c} - m\beta > 0. \tag{4.32}$$

By Lemma 4.4 with $r = 1$,

$$\mathsf{c} = \operatorname{codim} \mathcal{Z}_1(v) \geq m. \tag{4.33}$$

If $\alpha(v) = 1$, then $\beta < 1$, and therefore $\mathsf{c} - m\beta \geq m(1 - \beta) > 0$.

Now suppose that $\alpha(v) = \frac{d-1}{d-2}$. Then $\mathrm{wt}(v) + dm = 0$ and $\lambda \neq (d^k)$. In particular $d \mid |\lambda|$. Since equality in Lemma 4.4 cannot occur in the nonrectangular case, we have

$$\mathsf{c} \geq m + 1. \tag{4.34}$$

Moreover,

$$m \leq d - 2. \tag{4.35}$$

Indeed, if $m = d - 1$, then for every tableau occurring in the support of v,

$$-|\lambda| + dn_d + d(d-1) = 0.$$

Because $\lambda_d = 0$, one has $|\lambda| \leq d(d-1)$. The preceding identity therefore forces $n_d = 0$ and $|\lambda| = d(d-1)$, hence $\lambda = (d^{d-1})$, contrary to the hypothesis $\lambda \neq (d^k)$.

Combining (4.34) and (4.35), we get

$$\frac{\mathsf{c}}{m} \geq 1 + \frac{1}{m} \geq 1 + \frac{1}{d-2} = \frac{d-1}{d-2} = \alpha(v).$$

Since $\beta < \alpha(v)$, this implies (4.32). Thus the integral in (4.31) is finite in all cases, and by (4.29) this proves local integrability at the origin.

We have shown that $\|P(\xi)\|^{-\beta}$ is locally integrable near every zero of P. Since $\overline{\Omega} \cap Z(P)$ is compact and $\|P\|^{-\beta}$ is bounded away from $Z(P)$, a finite cover gives (4.24). This proves the lemma. ☐

4.4. **Triangular U-polynomials.** We next prove pointwise local integrability for the triangular polynomial maps which arise from (4.2).

Proposition 4.8 (Local integrability for triangular U-polynomials)**.** *Let $1 \leq \ell = \ell(\lambda) \leq \lambda_1$ be the smallest integer such that $\mu_\ell \geq 0$. Let*

$$Q(\xi) = \sum_{q=0}^{\ell} \frac{1}{q!} Y_\xi^q v_{\ell-q}, \qquad v_j \in V_j,$$

and assume that $Q \not\equiv 0$. Then for every zero ξ_0 of Q, and every $0 < \beta < \alpha_\lambda$, the function $\xi \to \|Q(\xi)\|^{-\beta}$ is locally integrable at ξ_0.

Proof. Fix a zero ξ_0 of Q. We shall verify the hypothesis of Lemma 4.6 for the singleton family $\{Q\}$ in a small neighborhood of ξ_0. Because U is abelian, for every $\xi \in \mathbb{R}^{d-1}$,

$$Q(\xi_0 + \xi) = \sum_{q=0}^{\ell} \frac{1}{q!} Y_\xi^q w_{\ell-q}, \tag{4.36}$$

where $w_j := p_j\Big(u_{\xi_0} \sum_{i=0}^{\ell} v_i\Big)$. Since $Q(\xi_0) = 0$, we have $w_\ell = 0$. Translation does not turn a nonzero polynomial into the zero polynomial. Hence some w_j is nonzero, and we may let $q \geq 1$ be the smallest integer such that $w_{\ell-q} \neq 0$. The terms of degree less than q in (4.36) vanish.

We claim that $q = m(w_{\ell-q})$. Indeed, $w_{\ell-q} \in V_{\ell-q}$, so its weight is $\mu_{\ell-q} = \mu_\ell - dq$. Since ℓ is the smallest index with $\mu_\ell \geq 0$, one has $0 \leq \mu_\ell < d$. Therefore the smallest integer m such that $\mu_{\ell-q} + dm \geq 0$ is exactly $m = q$.

Lemma 4.4, applied with $r = q$, now shows that $Y_\xi^q w_{\ell-q}$ is not identically zero. Therefore the first nonzero homogeneous Taylor term of Q at ξ_0 is

$$H_{Q,\xi_0}(\xi) = \frac{1}{q!} Y_\xi^q w_{\ell-q}. \tag{4.37}$$

Hence Lemma 4.7 applies to the weight homogeneous vector $w_{\ell-q}$, and

$$\mathrm{wt}(w_{\ell-q}) + dq = \mu_\ell.$$

It follows that the exponent $\alpha(w_{\ell-q})$ in Lemma 4.7 is precisely α_λ. The negative-moment conclusion established in the proof of Lemma 4.7 therefore gives, for every $0 < \beta < \alpha_\lambda$,

$$\int_{\|\xi\|<1} \|Y_\xi^q w_{\ell-q}\|^{-\beta}\, d\xi < \infty.$$

By (4.37), the same is true for the initial form H_{Q,ξ_0}. Since ξ_0 was arbitrary, the hypotheses of Lemma 4.6 hold for the singleton family $\{Q\}$ on any bounded semialgebraic neighborhood of ξ_0. Therefore for any $0 < \beta < \alpha_\lambda$, $\xi \to \|Q(\xi)\|^{-\beta}$ is locally integrable at ξ_0, as required. □

4.5. **Uniformity in the normalized family and proof of the sublevel estimate.** We now complete the proof of Proposition 4.1.

Proof of Proposition 4.1. Choose a bounded semialgebraic open set $B \subset \mathbb{R}^{d-1}$ such that $\overline{\Omega} \subset B$. It is enough to prove the asserted estimate with Ω replaced by B. Since all norms on a finite-dimensional vector space are equivalent, we may therefore work with Euclidean norms throughout the proof.

By homogeneity in v, it is enough to treat the normalized case

$$A_\ell(v) = 1. \tag{4.38}$$

Write $v_i := p_i(v)$ for $0 \le i \le \ell$. By (4.2),

$$P_v(\xi) := p_\ell(u_\xi v) = \sum_{q=0}^{\ell} \frac{1}{q!} Y_\xi^q v_{\ell-q}. \tag{4.39}$$

Thus P_v depends only on the tuple $(v_0, \dots, v_\ell)$.

Let

$$\mathcal{A}_\ell := \left\{ (v_0, \dots, v_\ell) \in V_0 \oplus \cdots \oplus V_\ell : \max_{0 \le i \le \ell} \|v_i\| = 1 \right\}.$$

This is a compact semialgebraic set. The map

$$\widetilde{P} : \mathcal{A}_\ell \times \mathbb{R}^{d-1} \to V_\ell, \qquad \widetilde{P}((v_0, \dots, v_\ell), \xi) = \sum_{q=0}^{\ell} \frac{1}{q!} Y_\xi^q v_{\ell-q},$$

is polynomial. Hence the collection

$$\mathcal{P} := \{P_v : A_\ell(v) = 1\}$$

is a compact semialgebraic family of polynomial maps $\mathbb{R}^{d-1} \to V_\ell$ of degree at most ℓ.

We claim that no element of $\mathcal{P}$ is identically zero. Let i_0 be the smallest index such that $v_{i_0} \neq 0$. The homogeneous term of degree $\ell - i_0$ in P_v is

$$\frac{1}{(\ell - i_0)!} Y_\xi^{\ell - i_0} v_{i_0}.$$

If $i_0 = \ell$, then this term is $v_\ell \neq 0$. Suppose $i_0 < \ell$. Then $v_{i_0} \in V_{i_0}$, and since ℓ is the smallest index with $\mu_\ell \geq 0$, one has $0 \leq \mu_\ell < d$. Hence $m(v_{i_0}) = \ell - i_0$. Applying Lemma 4.4 with $r = m(v_{i_0})$, we obtain

$$\operatorname{codim} \mathcal{Z}_{\ell - i_0}(v_{i_0}) \geq 1,$$

hence $Y_\xi^{\ell - i_0} v_{i_0} \not\equiv 0$. Thus $P_v \not\equiv 0$, and the claim follows.

Fix $0 < \beta < \alpha_\lambda$, and choose $\beta < \beta_1 < \alpha_\lambda$. By Proposition 4.8, every $P \in \mathcal{P}$ satisfies $\|P(\xi)\|^{-\beta_1} \in L^1_{\mathrm{loc}}(\xi_0)$ at every zero ξ_0 of P. Since $\overline{B} \cap Z(P)$ is compact and $\|P\|^{-\beta_1}$ is bounded away from the zero set, it follows that

$$\int_B \|P(\xi)\|^{-\beta_1} \, d\xi < \infty \qquad (P \in \mathcal{P}). \tag{4.40}$$

Identify $\mathcal{P}$ with its compact semialgebraic coefficient set $A \subset \mathbb{R}^N$. Then the evaluation map

$$Q : A \times B \to V_\ell, \qquad Q(P, \xi) := P(\xi),$$

is polynomial. Applying Lemma 4.5 to this family, with $s_1 = \beta_1/2$ and $s = \beta/2$, and using (4.40), we obtain

$$\sup_{P \in \mathcal{P}} \int_B \|P(\xi)\|^{-\beta} \, d\xi < \infty. \tag{4.41}$$

Consequently, Chebyshev's inequality gives

$$\mathrm{Leb}\{\xi \in B : \|P(\xi)\| \leq \varepsilon\} \ll_{\beta, B} \varepsilon^\beta \tag{4.42}$$

uniformly for all $P \in \mathcal{P}$ and all $0 < \varepsilon < 1$.

Now fix $\eta > 0$. If $\eta \geq \alpha_\lambda$, then $\varepsilon^{\alpha_\lambda - \eta} \geq 1$ for $0 < \varepsilon < 1$, so the required estimate is trivial after enlarging the constant. Assume $0 < \eta < \alpha_\lambda$, and choose $\alpha_\lambda - \eta < \beta < \alpha_\lambda$. Then (4.42) yields

$$\mathrm{Leb}\{\xi \in B : \|P_v(\xi)\| \leq \varepsilon\} \ll_{\eta, B} \varepsilon^{\alpha_\lambda - \eta}$$

for every v satisfying (4.38).

Finally, let $v \in V$ be arbitrary with $A_\ell(v) > 0$, and set $\widetilde{v} := A_\ell(v)^{-1} v$. Then $A_\ell(\widetilde{v}) = 1$ and $p_\ell(u_\xi \widetilde{v}) = A_\ell(v)^{-1} p_\ell(u_\xi v)$. Applying the normalized estimate to $\widetilde{v}$, we obtain

$$\begin{aligned} &\mathrm{Leb}\Big(\big\{\xi \in \Omega : \|p_\ell(u_\xi v)\| \leq \varepsilon A_\ell(v)\big\}\Big) \\ &\qquad = \mathrm{Leb}\Big(\big\{\xi \in \Omega : \|p_\ell(u_\xi \widetilde{v})\| \leq \varepsilon\big\}\Big) \ll_{\eta, \Omega} \varepsilon^{\alpha_\lambda - \eta}. \end{aligned}$$

This proves Proposition 4.1. $\square$

5. Weighted local estimates for the H-action on $\wedge^i \mathrm{M}_d(\mathbb{R})$

The purpose of this section is to convert the sublevel estimates of the previous section into local contraction inequalities for the representations that occur in the exterior powers of $\mathrm{M}_d(\mathbb{R})$. We work along the principal diagonal flow $a_t = (b_t, b_t)$ and average over compact subsets of its expanding horospherical subgroup. For generic irreducible summands, the sublevel estimates give contraction for a range of exponents slightly larger than 1. For the exceptional summands, the ordinary local heights exhibit a critical exponent at 1: one has contraction below this exponent and only bounded expansion above it. To pass beyond this obstruction, in the last subsection we introduce modified local height functions for the exceptional summands. These modified local heights recover the supercritical estimates needed in the global height inequalities of Section 6.

We retain the notation and conventions of the preceding section. Thus

$$V = \mathsf{S}_\lambda(\mathbb{R}^d), \qquad \lambda_d = 0,$$

where the Young diagram of λ is contained in the $d \times d$ box, and

$$V = \bigoplus_{j=0}^{\lambda_1} V_j, \qquad \mu_j = \mu_0 + dj = -|\lambda| + dj.$$

We assume that V is nontrivial, so $\mu_0 < 0$. In particular, $\mu_0 \in \mathbb{Z}$, $|\mu_0| \le d(d-1)$, and $d \mid \mu_0$ if and only if $d \mid |\lambda|$.

5.1. Contraction inequality for $\mathsf{S}_\lambda(\mathbb{R}^d)$.

Let α_λ be the exponent in Proposition 4.1, and define

$$\gamma_\lambda := \begin{cases} 1 & \text{if } \lambda = (d^k) \text{ for some } 1 \le k \le d-1 \\ 1 + d^{-2} & \text{otherwise.} \end{cases}$$

The main one-factor estimate is the following.

Proposition 5.1 (Contraction for $\|\cdot\|^{-\beta}$). *Let $0 < \beta < \gamma_\lambda$, and let $\Omega \subset \mathbb{R}^{d-1}$ be bounded and open. There exists $c = c(V, \beta, \Omega) > 0$ such that, for every $t \ge 0$ and every $0 \ne v \in V$,*

$$\int_\Omega \|b_t u_\xi v\|^{-\beta}\, d\xi \le c \exp\Big(-\frac{1}{4}\min\{\beta, \gamma_\lambda - \beta\} t\Big)\, \|v\|^{-\beta}.$$

We first record two elementary consequences of the polynomial structure from Section 4.

Lemma 5.2 (A distribution-function estimate). *Let $F : \Omega \to [0, \infty)$ be measurable, and suppose that, for some $a > 0$,*

$$\mathrm{Leb}\{\xi \in \Omega : F(\xi) \le \varepsilon\} \le C\varepsilon^a \qquad (0 < \varepsilon \le 1).$$

Then, for $s \ge 0$ and $\beta > 0$ with $\beta \ne a$,

$$\int_\Omega \max\{1, e^s F(\xi)\}^{-\beta}\, d\xi \ll_{C,a,\beta,\Omega} e^{-\min\{a,\beta\}s}.$$

Moreover, if $0 < \beta < a$ *and* $0 < \rho \le 1$*, then*

$$\int_{\{F\le\rho\}} \max\{e^{-s}, F(\xi)\}^{-\beta}\, d\xi \ll_{C,a,\beta,\Omega} \rho^{a-\beta}.$$

Proof. Decompose the range of F into the sets

$$\{F \le e^{-s}\}, \qquad \{e^{-s+j-1} < F \le e^{-s+j}\} \quad (1 \le j \le \lceil s \rceil),$$

and $\{F > 1\}$. The asserted bounds follow by summing the resulting geometric series. The truncated estimate is obtained in the same way, with the sum stopped at the first index for which $e^{-s+j} > \rho$. □

The next lemma supplies the weak higher-weight sublevel estimate needed in the borderline case $\mu_m = 0$.

Lemma 5.3 (Sublevel estimate at the next weight)**.** *Let* m *be the smallest index for which* $\mu_m \ge 0$*, and assume that* $\mu_m = 0$*. For every bounded open set* $\Omega \subset \mathbb{R}^{d-1}$*, there exists* $C = C(V, \Omega) > 0$ *such that for every* $v \in V$ *with* $A_m(v) > 0$ *and every* $0 < \rho < 1$*,*

$$\mathrm{Leb}\,\{\xi \in \Omega : \|p_{m+1}(u_\xi v)\| \le \rho A_m(v)\} \le C\rho^{1/(m+1)}$$

where $A_m(v) := \max_{0\le j\le m} \|p_j(v)\|$.

Proof. Since $\mu_m = 0$, one has $|\lambda| = dm$. Moreover, $m < \lambda_1$: otherwise $|\lambda| = d\lambda_1$, which, together with $\lambda_d = 0$, would force $\lambda = 0$. Thus V_{m+1} is defined.

We first show that no nonzero zero-weight vector is fixed by U. Suppose that $0 \ne w \in V_m$ and $Y_i w = 0$ for every $1 \le i \le d-1$. Let T be the largest semistandard tableau occurring in w. By Lemma 4.3, T has no i-free entry for any $i < d$. Hence every column of T contains an entry equal to d. Because $\lambda_d = 0$, each column has height at most $d-1$. A column of height h therefore contributes $(d-1) - (h-1) = d - h > 0$ to the $\log b_1$-weight. Thus T has positive weight, contradicting $T \in V_m$.

For $0 \le i \le m$, put $q = m - i$. If $0 \ne v_i \in V_i$, then Lemma 4.4, applied with $r = q = m(v_i)$, shows that $Y_\xi^q v_i \not\equiv 0$ when $q \ge 1$; the assertion is immediate for $q = 0$. We claim that

$$Y_\xi^{q+1} v_i \not\equiv 0.$$

Indeed, if this polynomial vanished identically, polarization would imply

$$Y_\eta Y_{\xi_1} \cdots Y_{\xi_q} v_i = 0$$

for all $\eta, \xi_1, \ldots, \xi_q$. Choose ξ such that $Y_\xi^q v_i \ne 0$, and take $\xi_1 = \cdots = \xi_q = \xi$. The resulting nonzero zero-weight vector is annihilated by every Y_η, contradicting the first paragraph.

It follows that the linear map

$$\bigoplus_{i=0}^{m} V_i \to \mathcal{P}_{\le m+1}(\mathbb{R}^{d-1}, V_{m+1})/\mathcal{P}_0(\mathbb{R}^{d-1}, V_{m+1}),$$

which sends $(v_0, \ldots, v_m)$ to the nonconstant part of

$$\sum_{i=0}^{m} \frac{1}{(m+1-i)!} Y_\xi^{m+1-i} v_i,$$

is injective. Here $\mathcal{P}_0$ denotes the constant polynomial maps; the homogeneous degrees $m+1-i$ are distinct, so no cancellation between different i's is possible.

Choose a closed ball B whose interior contains $\overline{\Omega}$. By equivalence of norms on the preceding finite-dimensional quotient, there exists $c_0 > 0$ such that

$$\sup_{\xi \in B} \|p_{m+1}(u_\xi v)\| \geq c_0 A_m(v)$$

for every v with $A_m(v) > 0$; the arbitrary constant term $p_{m+1}(v)$ is absorbed by the quotient norm. Choose $\xi_0 \in B$ at which the displayed supremum is attained, and a unit functional $\ell \in V_{m+1}^*$ satisfying $\ell(p_{m+1}(u_{\xi_0} v)) = \|p_{m+1}(u_{\xi_0} v)\|$. The scalar polynomial $P(\xi) := \ell(p_{m+1}(u_\xi v))$ has degree at most $m+1$ and satisfies

$$\sup_{\xi \in B} |P(\xi)| \geq c_0 A_m(v).$$

The multivariable Remez inequality [6] implies, for a scalar polynomial of degree at most $m+1$ on the fixed ball B,

$$\mathrm{Leb}\{\xi \in B : |P(\xi)| \leq \varepsilon \sup_B |P|\} \ll_{B,m} \varepsilon^{1/(m+1)}.$$

Since $\Omega \subset B$ and $|P(\xi)| \leq \|p_{m+1}(u_\xi v)\|$, this gives

$$\mathrm{Leb}\,\{\xi \in \Omega : \|p_{m+1}(u_\xi v)\| \leq \rho A_m(v)\} \ll_{V,\Omega} \rho^{1/(m+1)},$$

as required. □

We now prove Proposition 5.1.

Proof of Proposition 5.1. We may replace the given norm by an equivalent norm for which $\|v\| = \max_{0 \leq j \leq \lambda_1} \|p_j(v)\|$. This changes only the implicit constants. The assertion for $0 \leq t \leq 1$ follows from compactness, so we assume henceforth that $t \geq 1$. For $0 \leq q \leq \lambda_1$, set

$$A_q(v) := \max_{0 \leq j \leq q} \|p_j(v)\|.$$

Since Ω is bounded and the action of u_ξ is triangular with respect to the weight decomposition, there exists $C \geq 1$ such that

$$C^{-1}\|v\| \leq \|u_\xi v\| \leq C\|v\|, \qquad A_q(u_\xi v) \leq C A_q(v) \tag{5.1}$$

for every $\xi \in \Omega$, every q, and every $v \in V$.

Let

$$m := \min\{j : \mu_j \geq 0\} = \left\lceil \frac{|\mu_0|}{d} \right\rceil.$$

We distinguish two cases.

Case 1: $A_m(v) < (2C^2)^{-1}\|v\|$.

If $m = \lambda_1$, this case is empty. Otherwise, (5.1) implies

$$\max_{m+1 \le j \le \lambda_1} \|p_j(u_\xi v)\| \ge C^{-1}\|v\|$$

after increasing C harmlessly. Since $\mu_{m+1} \ge d$,

$$\|b_t u_\xi v\| \ge C^{-1} e^{\mu_{m+1} t}\|v\| \ge C^{-1} e^{dt}\|v\|.$$

Consequently,

$$\int_\Omega \|b_t u_\xi v\|^{-\beta}\, d\xi \ll_{\beta,\Omega} e^{-d\beta t}\|v\|^{-\beta},$$

which is stronger than the required estimate.

Case 2: $A_m(v) \ge (2C^2)^{-1}\|v\|$.

The triangular formula (4.2) and the boundedness of Ω give a constant $K \ge 1$ such that

$$\|p_j(u_\xi v) - p_j(v)\| \le K \max_{0 \le i < j} \|p_i(v)\| \qquad (1 \le j \le m,\ \xi \in \Omega).$$

Choose an index at which $A_m(v)$ is attained. If the maximum of the preceding coordinates is larger than $(2K)^{-1}$ times the current coordinate, move to an earlier index where that maximum is attained, and repeat. The process stops after at most m steps. Its final index ℓ satisfies

$$\|p_\ell(v)\| \ge (2K)^{-m} A_m(v) \quad \text{and} \quad \max_{0 \le i < \ell} \|p_i(v)\| \le (2K)^{-1}\|p_\ell(v)\|. \tag{5.2}$$

Consequently, for $c_0 := 2^{-1}(2K)^{-m}$,

$$\|p_\ell(u_\xi v)\| \ge \frac{1}{2}\|p_\ell(v)\| \ge c_0 A_m(v) \qquad (\xi \in \Omega). \tag{5.3}$$

Put

$$A := A_m(v), \qquad F(\xi) := \frac{\|p_m(u_\xi v)\|}{A}, \qquad s := d(m-\ell)t.$$

By Proposition 4.1, for every $\eta > 0$,

$$\mathrm{Leb}\{\xi \in \Omega : F(\xi) \le \varepsilon\} \ll_{\eta,V,\Omega} \varepsilon^{\alpha_\lambda - \eta} \qquad (0 < \varepsilon < 1). \tag{5.4}$$

Moreover, (5.3) gives

$$\|b_t u_\xi v\| \gg A e^{\mu_\ell t} \max\{1, e^s F(\xi)\}. \tag{5.5}$$

We first suppose that $\mu_m > 0$, equivalently $d \nmid |\lambda|$. In this case $\alpha_\lambda = 1$ and $\gamma_\lambda = 1 + d^{-2}$. Set

$$\gamma := \frac{dm}{|\mu_0|} = 1 + \frac{\mu_m}{|\mu_0|}.$$

Since μ_m is a positive integer and $|\mu_0| \le d(d-1)$,

$$\gamma \ge 1 + \frac{1}{d(d-1)} > 1 + \frac{1}{d^2} = \gamma_\lambda.$$

If $0 < \beta < 1$, choose a with $\beta < a < 1$ in (5.4). Lemma 5.2 and (5.5) yield

$$\int_\Omega \|b_t u_\xi v\|^{-\beta}\, d\xi \ll e^{-\beta\mu_m t} A^{-\beta} \ll e^{-\beta t}\|v\|^{-\beta}.$$

Now let $1 \le \beta < \gamma_\lambda$. Choose $a = 1 - \eta$ so close to 1 that

$$\gamma a - \beta \ge \frac{1}{2}(\gamma - \beta).$$

If $\ell = m$, (5.5) already gives decay at rate $\beta\mu_m \ge \beta$. If $\ell < m$, then

$$\frac{d(m-\ell)}{|\mu_\ell|} = \frac{d(m-\ell)}{|\mu_0| - d\ell} \ge \gamma.$$

Applying Lemma 5.2 with exponent a therefore gives a decay exponent at least

$$|\mu_\ell|(\gamma - \beta) \ge \frac{1}{2}(\gamma - \beta) \ge \frac{1}{2}(\gamma_\lambda - \beta).$$

This proves the proposition when $\mu_m > 0$.

It remains to consider $\mu_m = 0$. Then

$$\alpha_\lambda = \begin{cases} 1, & \lambda = (d^k), \\ \dfrac{d-1}{d-2}, & \lambda \ne (d^k), \end{cases} \qquad \gamma_\lambda \le \alpha_\lambda.$$

Choose

$$a := \alpha_\lambda - \tfrac{1}{10}(\alpha_\lambda - \beta),$$

so that $a > \beta$. For $0 < \rho < 1$, define

$$D^+(\rho) := \left\{\xi \in \Omega : \max_{m+1 \le j \le \lambda_1} \|p_j(u_\xi v)\| \le \rho A\right\}$$

and

$$D(\rho) := \left\{\xi \in \Omega : \max_{m \le j \le \lambda_1} \|p_j(u_\xi v)\| \le \rho A\right\}.$$

Let $0 < \rho_1 < \rho_2 \le 1$. On $\Omega - D^+(\rho_1)$,

$$\|b_t u_\xi v\| \ge e^{dt}\rho_1 A,$$

and hence

$$\int_{\Omega - D^+(\rho_1)} \|b_t u_\xi v\|^{-\beta}\, d\xi \ll e^{-d\beta t}\rho_1^{-\beta} A^{-\beta}. \tag{5.6}$$

If $\xi \in D^+(\rho_1) - D(\rho_2)$, then $\|p_m(u_\xi v)\| > \rho_2 A$. By Lemma 5.3,

$$\mathrm{Leb}(D^+(\rho_1)) \ll \rho_1^{1/(m+1)},$$

so

$$\int_{D^+(\rho_1) - D(\rho_2)} \|b_t u_\xi v\|^{-\beta}\, d\xi \ll \rho_1^{1/(m+1)} \rho_2^{-\beta} A^{-\beta}. \tag{5.7}$$

Finally, on $D(\rho_2)$ one has $F \le \rho_2$. Since $\mu_\ell = -d(m-\ell)$, (5.5) becomes

$$\|b_t u_\xi v\| \gg A \max\{e^{-L}, F(\xi)\}.$$

The truncated part of Lemma 5.2, together with (5.4), yields

$$\int_{D(\rho_2)} \|b_t u_\xi v\|^{-\beta}\, d\xi \ll \rho_2^{a-\beta} A^{-\beta}. \tag{5.8}$$

Take

$$\rho_1 = e^{-(d-1)t}, \qquad \rho_2 = e^{-\frac{d-1}{2(m+1)}t}.$$

Since $m \le d-1$, the three terms in (5.6)–(5.8) are bounded respectively by

$$e^{-\beta t}A^{-\beta}, \qquad e^{-\frac{d-1}{m+1}(1-\beta/2)t}A^{-\beta}, \qquad e^{-\frac{d-1}{2(m+1)}(a-\beta)t}A^{-\beta}.$$

Here $1-\beta/2 \ge 4/9$, while

$$a-\beta = \tfrac{9}{10}(\alpha_\lambda - \beta) \ge \tfrac{9}{10}(\gamma_\lambda - \beta).$$

Moreover, $m+1 \le d$, and hence

$$\frac{d-1}{m+1}\left(1-\frac{\beta}{2}\right) \ge \frac{8}{27}, \qquad \frac{d-1}{2(m+1)}(a-\beta) \ge \tfrac{3}{10}(\gamma_\lambda - \beta).$$

Because $A \asymp \|v\|$ in Case 2, the three contributions therefore imply

$$\int_\Omega \|b_t u_\xi v\|^{-\beta}\, d\xi \ll \exp\Bigl(-\frac{1}{4}\min\{\beta, \gamma_\lambda - \beta\} t\Bigr)\, \|v\|^{-\beta}.$$

This completes the proof. $\square$

Corollary 5.4 (Bounded expansion for $\|\cdot\|^{-\beta}$)**.** *Let $\beta > \gamma_\lambda$, and let $\Omega \subset \mathbb{R}^{d-1}$ be bounded and open. There exists $c = c(V, \beta, \Omega) > 0$ such that*

$$\int_\Omega \|b_t u_\xi v\|^{-\beta}\, d\xi \le c e^{d^2(\beta-\gamma_\lambda)t}\|v\|^{-\beta}$$

for every $t \ge 0$ and every $0 \ne v \in V$.

Proof. Uniformly for $\xi \in \Omega$,

$$\|b_t u_\xi v\| \gg e^{-d(d-1)t}\|v\|,$$

because the smallest b_t-weight is at least $-d(d-1)$. Put $x = \beta - \gamma_\lambda > 0$ and

$$\delta := \min\left\{\frac{x}{d-1}, \frac{\gamma_\lambda}{2}\right\}, \qquad \beta_0 := \gamma_\lambda - \delta.$$

Then $0 < \beta_0 < \gamma_\lambda$, so Proposition 5.1 gives a uniform bound for the β_0-moment. Therefore

$$\begin{aligned}\int_\Omega \|b_t u_\xi v\|^{-\beta}\, d\xi &\ll e^{d(d-1)(\beta-\beta_0)t}\|v\|^{-(\beta-\beta_0)} \int_\Omega \|b_t u_\xi v\|^{-\beta_0}\, d\xi \\ &\ll e^{d(d-1)(x+\delta)t}\|v\|^{-\beta}.\end{aligned}$$

The definition of δ gives $d(d-1)(x+\delta) \le d^2 x$, proving the claim. $\square$

5.2. **One-factor estimates along the principal ray.** We now specialize to the representations occurring in

$$\mathcal{V}_{\text{all}} := \bigoplus_{1 \le i \le d^2-1} \wedge^i \mathrm{M}_d(\mathbb{R}),$$

viewed as a representation of $H = \mathrm{SL}_d(\mathbb{R}) \times \mathrm{SL}_d(\mathbb{R})$. Recalling (3.3), set

$$\mathcal{V}_{\text{gen}} := \bigoplus_{1 \le i \le d^2-1,\ d \nmid i} \wedge^i \mathrm{M}_d(\mathbb{R}) \oplus \mathcal{M}_0.$$

Lemma 5.5 (Contraction on generic $\mathrm{SL}_d(\mathbb{R})$-types). *Let $V_1 \otimes V_2 \subset \mathcal{V}_{\mathrm{gen}}$ be an irreducible H-subrepresentation, let $m \in \{1,2\}$, and let $\Omega \subset \mathbb{R}^{d-1}$ be bounded and open. For every $\frac{1}{2} \le \beta \le 1 + \frac{1}{2d^2}$, there exists $c = c(V_m, \beta, \Omega) > 0$ such that*

$$\int_\Omega \|b_t u_\xi v\|^{-\beta}\, d\xi \le c e^{-t/(8d^2)} \|v\|^{-\beta}$$

for every $t \ge 0$ and every nonzero $v \in V_i$.

Proof. By the definition of $\mathcal{V}_{\mathrm{gen}}$, the reduced partition corresponding to V_i is not of the form (d^k). Hence $\gamma_\lambda = 1 + d^{-2}$. Since $\min\left\{\beta, 1 + \frac{1}{d^2} - \beta\right\} \ge \frac{1}{2d^2}$, Proposition 5.1, applied to b_t, gives

$$\int_\Omega \|b_t u_\xi v\|^{-\beta}\, d\xi \ll e^{-t/(8d^2)} \|v\|^{-\beta},$$

which is the asserted estimate. □

The same estimate holds for every finite direct sum of generic types. Indeed, use the maximum norm, choose a component of maximal norm, and apply the preceding lemma to that invariant component. In addition, Hölder's inequality gives

$$\int_\Omega \|b_t u_\xi w\|^{-q}\, d\xi \ll_{q,\Omega} \|w\|^{-q} \qquad \left(0 \le q \le 1 + \frac{1}{2d^2}\right) \tag{5.9}$$

for every such finite direct sum.

Lemma 5.6 (Estimates for $\mathcal{M}_{k,1}$ and $\mathcal{M}_{k,2}$). *Let $1 \le k \le d-1$, $m \in \{1,2\}$, and let $\Omega \subset \mathbb{R}^{d-1}$ be bounded and open. Fix $0 < \theta \le (2d)^{-6}$. For every β satisfying $\frac{1}{2} \le \beta \le 1 - \frac{\theta}{10}$ or $\beta \ge 1 + \frac{\theta}{10}$, there exists $c = c(\beta, \theta, \Omega) > 0$ such that, for every $t \ge 0$ and every nonzero vector in the nontrivial H_i-module $\mathcal{M}_{k,i}$,*

$$\int_\Omega \|b_t u_\xi v\|^{-\beta}\, d\xi \le \begin{cases} c e^{-\frac{1}{5}(1-\beta)t} \|v\|^{-\beta}, & \frac{1}{2} \le \beta \le 1 - \frac{\theta}{10}, \\ c e^{2d^2(\beta-1)t} \|v\|^{-\beta}, & \beta \ge 1 + \frac{\theta}{10}. \end{cases}$$

Proof. As an $H_m \simeq \mathrm{SL}_d(\mathbb{R})$-module, $\mathcal{M}_{k,m} \simeq \mathsf{S}_{(d^k)}(\mathbb{R}^d)$, so $\gamma_\lambda = 1$. For $\beta \le 1 - \theta/10$, Proposition 5.1 gives the first estimate. For $\beta \ge 1 + \theta/10$, apply Corollary 5.4. □

5.3. **Local estimates for $H = \mathrm{SL}_d(\mathbb{R}) \times \mathrm{SL}_d(\mathbb{R})$.** We now pass from the one-factor estimates to estimates for the full left–right action, applying them successively in the two horospherical variables and using the tensor-product structure of the H-summands.

For $\xi = (\xi_1, \xi_2) \in \mathbb{R}^{d-1} \times \mathbb{R}^{d-1}$, put

$$n_\xi := \left(\begin{pmatrix} \mathrm{Id}_{d-1} & 0 \\ \xi_1 & 1 \end{pmatrix}, \begin{pmatrix} \mathrm{Id}_{d-1} & 0 \\ \xi_2 & 1 \end{pmatrix} \right) \in H.$$

We write $u_1(\xi_1) = n_{(\xi_1,0)}$ and $u_2(\xi_2) = n_{(0,\xi_2)}$, so that $n_\xi = u_1(\xi_1) u_2(\xi_2)$. Let

$$U_m := \{u_m(\xi_m) : \xi_m \in \mathbb{R}^{d-1}\}, \qquad N := U_1 U_2.$$

Write

$$B_{U_m}(r) := \{u_m(\xi_m) \in U_m : \|\xi_m\| \le r\}.$$

Using the maximum norm $\|\xi\| = \max(\|\xi_1\|, \|\xi_2\|)$, we have $B_N(r) = B_{U_1}(r)B_{U_2}(r)$.

Proposition 5.7. *Let $\Omega \subset \mathbb{R}^{d-1} \times \mathbb{R}^{d-1}$ be bounded and open. For every $\frac{1}{2} \le \beta \le 1 + \frac{1}{2d^2}$, there exists $c > 0$ such that for every $t \ge 0$ and every nonzero $v \in \mathcal{V}_{\mathrm{gen}}$,*

$$\int_\Omega \|a_t n_\xi v\|^{-\beta}\, d\xi \le c e^{-t/(5d^2)} \|v\|^{-\beta}. \tag{5.10}$$

Moreover, for $0 \le \beta \le 1 + \frac{1}{2d^2}$,

$$\int_\Omega \|a_t n_\xi v\|^{-\beta}\, d\xi \ll_{\beta,\Omega} \|v\|^{-\beta}. \tag{5.11}$$

Proof. Enlarge Ω to a product $\Omega_1 \times \Omega_2$. Choose the maximum norm associated with the finite decomposition of $\mathcal{V}_{\mathrm{gen}}$ into irreducible H-modules. For an arbitrary nonzero v, choose an irreducible component v' with $\|v'\| = \|v\|$. Since the summands are H-invariant, $\|a_t n_\xi v\| \ge \|a_t n_\xi v'\|$. It therefore suffices to treat $v \in V_1 \otimes V_2$ in a fixed irreducible summand.

Viewed as an H_1-module, $V_1 \otimes V_2$ is a finite direct sum of copies of V_1; similarly for H_2. Thus the direct-sum extension of Lemma 5.5, followed by Fubini's theorem, gives

$$\begin{aligned}
\int_{\Omega_2} \int_{\Omega_1} &\|(b_t u_{\xi_1}, b_t u_{\xi_2})v\|^{-\beta}\, d\xi_1 d\xi_2 \\
&\ll e^{-t/(10d^2)} \int_{\Omega_2} \|(e, b_t u_{\xi_2})v\|^{-\beta}\, d\xi_2 \\
&\ll e^{-t/(5d^2)} \|v\|^{-\beta}.
\end{aligned}$$

This proves (5.10). The bounded estimate for $\beta < 1/2$ follows from the case $\beta = 1/2$ by Hölder's inequality; the remaining range follows directly from the contraction estimate. □

Proposition 5.8 (Bounded expansion on $\mathcal{M}_{k,1} \oplus \mathcal{M}_{k,2}$)**.** *Let $1 \le k \le d-1$, $0 < \theta < (2d)^{-6}$, and let $\Omega \subset \mathbb{R}^{d-1} \times \mathbb{R}^{d-1}$ be bounded and open. For every β satisfying $\frac{1}{2} \le \beta \le 1 - \frac{\theta}{10}$ or $\beta \ge 1 + \frac{\theta}{10}$, there exists $c = c(\beta, \theta, \Omega) > 0$ such that, for every $t \ge 0$ and every $0 \ne v \in \mathcal{M}_{k,1} \oplus \mathcal{M}_{k,2}$,*

$$\int_\Omega \|a_t n_\xi v\|^{-\beta}\, d\xi \le \begin{cases} c e^{-\frac{1}{5}(1-\beta)t} \|v\|^{-\beta}, & \frac{1}{2} \le \beta \le 1 - \frac{\theta}{10}, \\ c e^{2d^2(\beta-1)t} \|v\|^{-\beta}, & \beta \ge 1 + \frac{\theta}{10}. \end{cases}$$

Proof. Use the maximum norm on $\mathcal{M}_{k,1} \oplus \mathcal{M}_{k,2}$, and write $v = v_1 + v_2$. Choose $m \in \{1, 2\}$ so that $\|v_m\| = \|v\|$. Since the two summands are H-invariant,

$$\|a_t n_\xi v\| \ge \|a_t n_\xi v_m\|.$$

If $m = 1$, the second factor acts trivially and

$$a_t n_\xi v_1 = b_t u_{\xi_1} v_1;$$

the case $m = 2$ is symmetric. The result follows from Lemma 5.6, after integrating over the irrelevant variable. □

5.4. **Modified local height functions and local contraction.** At a critical degree kd, the ordinary local height $\|\cdot\|^{-\beta}$ has critical exponent 1 because of the two exceptional summands $\mathcal{M}_{k,1}$ and $\mathcal{M}_{k,2}$. To obtain contraction in a supercritical range of exponents, we modify the local height by penalizing not only small vectors, but also proximity to either exceptional summand.

Definition 5.9 (Modified local height)**.** Fix $1 \le k \le d-1$ and $0 < \theta < (2d)^{-6}$. For $v \notin \mathcal{M}_{k,1} \cup \mathcal{M}_{k,2}$, define

$$\phi_{kd,\theta}(v) := \|v\|^{-1+4\theta} \|v - \pi_{k,1}(v)\|^{-2\theta} \|v - \pi_{k,2}(v)\|^{-2\theta}.$$

For $v \in \mathcal{M}_{k,1} \cup \mathcal{M}_{k,2}$, set $\phi_{kd,\theta}(v) = \infty$. We write $\phi_{kd} = \phi_{kd,\theta}$ when θ is fixed.

Choose the maximum norm associated with $\mathcal{M}_{k,0} \oplus \mathcal{M}_{k,1} \oplus \mathcal{M}_{k,2}$, and write $v_m = \pi_{k,m}(v)$. With the convention $0^{-q} = \infty$, define

$$F_{00}(v) := \|v_0\|^{-1},$$

$$F_{12}(v) := \|v_1\|^{-1+2\theta} \|v_2\|^{-2\theta}, \qquad F_{21}(v) := \|v_2\|^{-1+2\theta} \|v_1\|^{-2\theta},$$

and

$$F_{10}(v) := \|v_1\|^{-1+2\theta} \|v_0\|^{-2\theta}, \qquad F_{20}(v) := \|v_2\|^{-1+2\theta} \|v_0\|^{-2\theta}.$$

A direct comparison of the three component norms gives

$$\phi_{kd}(v) = \min\{F_{00}(v), F_{12}(v), F_{21}(v), F_{10}(v), F_{20}(v)\}. \tag{5.12}$$

Equivalently,

$$\phi_{kd}(v) = \begin{cases} \|v_0\|^{-1}, & \|v\| = \|v_0\|, \\ \|v_1\|^{-1+2\theta} \max\{\|v_0\|, \|v_2\|\}^{-2\theta}, & \|v\| = \|v_1\|, \\ \|v_2\|^{-1+2\theta} \max\{\|v_0\|, \|v_1\|\}^{-2\theta}, & \|v\| = \|v_2\|. \end{cases}$$

Proposition 5.10 (Contraction for ϕ_{kd})**.** *Let $1 \le k \le d-1$, $0 < \theta < (2d)^{-6}$, and $1 \le \beta \le 1 + \frac{\theta}{4d^2}$. There exists $c = c(\beta, \theta) > 0$ such that, for every $t \ge 0$ and every $v \notin \mathcal{M}_{k,1} \cup \mathcal{M}_{k,2}$,*

$$\int_{B_N(1)} \phi_{kd}(a_t n v)^\beta \, dn \le c e^{-\theta t/(20d^2)} \phi_{kd}(v)^\beta.$$

The rest of this section is devoted to the proof of this proposition. We prove the five estimates corresponding to (5.12).

Lemma 5.11. *Let* $1 \leq \beta \leq 1 + \frac{1}{2d^2}$*. If* $v_0 \neq 0$*, then*

$$\int_{B_N(1)} \|\pi_{k,0}(a_t n v)\|^{-\beta}\, dn \ll e^{-t/(5d^2)} \|v_0\|^{-\beta}$$

for every $t \geq 0$*.*

Proof. The projections are H-equivariant, and $\mathcal{M}_{k,0} \subset \mathcal{V}_{\text{gen}}$. Apply Proposition 5.7 to v_0. □

Lemma 5.12. *Let* $1 \leq \beta \leq 1 + \frac{\theta}{4d^2}$*, and put*

$$q := (1-2\theta)\beta, \qquad r := 2\theta\beta.$$

If $v_1, v_2 \neq 0$*, then*

$$\int_{B_N(1)} \|\pi_{k,1}(a_t n v)\|^{-q} \|\pi_{k,2}(a_t n v)\|^{-r}\, dn \ll e^{-\theta t/(20d^2)} \|v_1\|^{-q} \|v_2\|^{-r}.$$

The same estimate holds with 1 *and* 2 *interchanged.*

Proof. We have $q \geq \frac{1}{2}$, $r < \frac{1}{2}$, and

$$1 - q = 1 - (1-2\theta)\beta \geq 2\theta - \frac{\theta}{4d^2} > \theta.$$

The first factor depends only on u_1, and the second only on u_2. Lemma 5.6 therefore gives

$$\int_{B_{U_1}(1)} \|b_t u_1 v_1\|^{-q}\, du_1 \ll e^{-(1-q)t/5} \|v_1\|^{-q}.$$

For the exponent $r < 1/2$, Hölder's inequality and the estimate at exponent 1/2 give

$$\int_{B_{U_2}(1)} \|b_t u_2 v_2\|^{-r}\, du_2 \ll \|v_2\|^{-r}.$$

Fubini's theorem completes the proof. □

Lemma 5.13. *Let* $m \in \{1,2\}$*,* $1 \leq \beta \leq 1 + \frac{\theta}{4d^2}$*, and put*

$$q_0 := 2\theta\beta, \qquad q_1 := (1-2\theta)\beta.$$

If $v_0, v_m \neq 0$*, then*

$$\int_{B_N(1)} \|\pi_{k,0}(a_t n v)\|^{-q_0} \|\pi_{k,m}(a_t n v)\|^{-q_1}\, dn \ll e^{-\theta t/(20d^2)} \|v_0\|^{-q_0} \|v_m\|^{-q_1}.$$

Proof. We treat $m = 1$; the other case is symmetric. Put

$$s := 1 + \frac{1}{2d^2}, \qquad a := \frac{q_0}{s}, \qquad r := \frac{q_1}{1-a}.$$

Then $0 < a < 1$, and

$$\begin{aligned} 1 - a - q_1 &= 1 - \beta + 2\theta\beta\left(1 - \frac{1}{s}\right) \\ &= 1 - \beta + \frac{2\theta\beta}{2d^2+1} \geq -\frac{\theta}{4d^2} + \frac{2\theta}{2d^2+1} \geq \frac{\theta}{2d^2}. \end{aligned}$$

Consequently $\frac{1}{2} < r < 1$, and

$$(1-a)(1-r) = 1 - a - q_1 \geq \frac{\theta}{2d^2}. \tag{5.13}$$

Write $g_j = b_t u_j$. For fixed u_2, Hölder's inequality in the u_1-variable gives

$$\begin{aligned}&\int_{B_{U_1}(1)} \|(g_1, g_2)v_0\|^{-q_0} \|g_1 v_1\|^{-q_1} \, du_1 \\ &\quad \leq \left(\int_{B_{U_1}(1)} \|(g_1, g_2)v_0\|^{-s} \, du_1 \right)^a \left(\int_{B_{U_1}(1)} \|g_1 v_1\|^{-r} \, du_1 \right)^{1-a}.\end{aligned}$$

Every irreducible H_1-type occurring in $\mathcal{M}_{k,0}$ is nontrivial and generic. Indeed, a trivial reduced partition would give the summand $\mathcal{M}_{k,2}$. If the reduced partition were (d^j) with $j \geq 1$, the $d \times d$ box condition would force the original partition itself to be (d^j). Since its size is kd, one would have $j = k$, giving the summand $\mathcal{M}_{k,1}$. Both possibilities are excluded from $\mathcal{M}_{k,0}$. Hence the direct-sum form of Lemma 5.5 gives

$$\int_{B_{U_1}(1)} \|(g_1, g_2)v_0\|^{-s} \, du_1 \ll e^{-t/(10d^2)} \|(e, g_2)v_0\|^{-s}.$$

Proposition 5.1, applied along b_t, gives

$$\int_{B_{U_1}(1)} \|g_1 v_1\|^{-r} \, du_1 \ll e^{-(1-r)t/5} \|v_1\|^{-r}.$$

Using (5.13) and then integrating in u_2, we obtain

$$\begin{aligned}&\int_{B_N(1)} \|a_t n v_0\|^{-q_0} \|a_t n v_1\|^{-q_1} \, dn \\ &\quad \ll e^{-\theta t/(20d^2)} \|v_1\|^{-q_1} \int_{B_{U_2}(1)} \|(e, g_2)v_0\|^{-q_0} \, du_2.\end{aligned}$$

Finally, $q_0 < 1/2$, so (5.9), applied to the H_2-action on $\mathcal{M}_{k,0}$, bounds the last integral by $\|v_0\|^{-q_0}$. □

Proof of Proposition 5.10. For each $v \notin \mathcal{M}_{k,1} \cup \mathcal{M}_{k,2}$, choose one of the five functions in (5.12), say F, such that

$$\phi_{kd}(v) = F(v).$$

This chosen value is finite, so every component appearing with a negative power in F is nonzero. Since $\phi_{kd}(w) \leq F(w)$ for every w,

$$\int_{B_N(1)} \phi_{kd}(a_t n v)^\beta \, dn \leq \int_{B_N(1)} F(a_t n v)^\beta \, dn.$$

If $F = F_{00}$, apply Lemma 5.11; if $F = F_{12}$ or F_{21}, apply Lemma 5.12; and if $F = F_{10}$ or F_{20}, apply Lemma 5.13. In every case the resulting right-hand side is at most

$$Ce^{-\theta t/(20d^2)} F(v)^\beta = Ce^{-\theta t/(20d^2)} \phi_{kd}(v)^\beta.$$

This proves the proposition. □

6. The global Margulis inequality for the modified height

Throughout this section the lattice is arbitrary. The goal is to prove one of the main technical ingredients of the paper: a global Margulis inequality for the modified height outside the exceptional set (Theorem 6.9). This estimate globalizes the local estimates of Section 5 and will later be combined with avoidance and iteration to obtain the uniform $L^{1+\theta}$-bound.

The estimate is not a pure contraction inequality: it has a logarithmic loss, reflecting the possible number of competing short rational subspaces, and an additive exponential error, which will be absorbed by the choice of step sizes in the iteration. When several rational subspaces become short simultaneously, the usual Plücker inequality [13, Lemma 5.6] controls their covolumes but not their projections to the exceptional summands. The Mother Inequality of Benoist–Quint supplies this missing control for intersections and sums.

Put $N = d^2$, and fix

$$0 < \theta < (2d)^{-6}.$$

For $1 \le i \le N-1$, define

$$\theta_i := 10^{i-N}\theta \tag{6.1}$$

and

$$\beta_i := \left(1 - \frac{\theta_i}{8d^2}\right)^{-1}. \tag{6.2}$$

We also set

$$\tau_i := (1 + \kappa i(N-i)\theta_1)^{-1}, \qquad \kappa := \frac{1}{16d^2(N-1)}. \tag{6.3}$$

Finally, put

$$\vartheta := \frac{\theta_1}{100d}.$$

The numerical choices in this paragraph are bookkeeping devices. Their only purpose is to create a hierarchy of small exponents with enough room for the local contraction estimates, the intersection–sum inequalities, and the iteration argument. The parameters $\beta_i > 1$ will be used for the modified local heights; the parameters $\tau_i < 1$ will be used for the ordinary height $\bar{\alpha}$, where the concavity in the index i is needed when two short rational subspaces are replaced by their intersection and sum. The rapidly increasing sequence $\theta_1 \ll \theta_2 \ll \cdots \ll \theta_{N-1}$ is chosen only to make these concavity inequalities strict. Along the principal ray, no additional cone restriction is needed to apply simultaneously the generic and exceptional local estimates from Section 5.

The choices above have three elementary consequences. First,

$$1 < \beta_i \le 1 + \frac{1}{2d^2} \qquad (1 \le i \le N-1), \tag{6.4}$$

and, if $i = kd$, then

$$1 \le \beta_i \le 1 + \frac{\theta_i}{4d^2}. \tag{6.5}$$

Indeed, $(1-x)^{-1} \le 1+2x$ for the values of $x = \theta_i/(8d^2)$ under consideration. Thus β_{kd} lies in the range of Proposition 5.10.

Second,

$$\frac{1}{2} < \tau_i < 1, \qquad 1-\tau_i \ge \frac{\vartheta}{10} \quad \text{whenever } d \mid i. \tag{6.6}$$

For the second assertion, write $x_i = \kappa i(N-i)\theta_1$. If $d \mid i$, then

$$x_i \ge \kappa d(N-d)\theta_1 = \frac{\theta_1}{16(d+1)}.$$

Since $x_i < 1$, it follows that $1-\tau_i = x_i/(1+x_i) \ge x_i/2 \ge \vartheta/10$.

Third,

$$\beta_i^{-1} < \tau_i \qquad (1 \le i \le N-1). \tag{6.7}$$

Indeed, $\beta_i^{-1} = 1-\theta_i/(8d^2)$, whereas

$$\kappa i(N-i)\theta_1 \le \frac{\theta_i}{16d^2};$$

the latter inequality is immediate for $i=1$, and for $i \ge 2$ it follows from $i(N-i) \le i(N-1) \le 2^{i-1}(N-1) \le 10^{i-1}(N-1)$. Therefore

$$\tau_i = \frac{1}{1+x_i} \ge 1-x_i > 1-\frac{\theta_i}{8d^2} = \beta_i^{-1}.$$

For later endpoint conventions, put

$$\alpha_0^+(\Delta) = \alpha_N^+(\Delta) := 1, \qquad \alpha_i^+(\Delta) := \max\{1, \alpha_i(\Delta)\} \quad (1 \le i \le N-1).$$

6.1. **The global and modified heights.** Define

$$\bar{\alpha}_\theta(\Delta) := \max_{1\le i\le N-1} \alpha_i^+(\Delta)^{\tau_i}. \tag{6.8}$$

We also introduce a second auxiliary height, $\bar{\alpha}'_\theta$, used only for bounded-expansion estimates and for absorbing the error terms coming from pairs of short subspaces:

$$\bar{\alpha}'_\theta(\Delta) := \sum_{i=1}^{N-1} \bar{\alpha}'_{\theta,i}(\Delta). \tag{6.9}$$

where, putting

$$\chi := \frac{\theta_1}{100d^2}, \qquad \sigma := \frac{3}{4},$$

we define

$$\bar{\alpha}'_{\theta,i}(\Delta) := \alpha_i^+(\Delta)^{1+\chi} + \sum_{1\le j\le \min\{i,N-i\}} \left(\alpha_{i-j}^+(\Delta)\alpha_{i+j}^+(\Delta)\right)^\sigma. \tag{6.10}$$

Here we use the endpoint convention $\alpha_0^+(\Delta) = \alpha_N^+(\Delta) = 1$. We usually write $\bar{\alpha}, \bar{\alpha}'$ when θ is fixed. Since $\tau_i \ge 1-C_d\theta_1$, one has

$$\max\{1,\alpha(\Delta)\}^{1-C_d\theta_1} \le \bar{\alpha}(\Delta) \le \max\{1,\alpha(\Delta)\}, \tag{6.11}$$

and

$$\max\{1,\alpha(\Delta)\}^{1+\chi} \le \bar{\alpha}'(\Delta) \ll_d \max\{1,\alpha(\Delta)\}^2. \tag{6.12}$$

Here and below, C_d denotes a positive constant depending only on d, whose value may change from line to line. Unsubscripted implied constants may also depend on the fixed parameter θ.

All height and quasi-null definitions continue to use the fixed Euclidean norm from Section 3. When we classify a component as maximal, we use the maximum of the three component norms associated with $\mathcal{M}_{k,0}\oplus\mathcal{M}_{k,1}\oplus\mathcal{M}_{k,2}$; this auxiliary maximum is uniformly equivalent to the Euclidean norm.

Definition 6.1 (Auxiliary weighted modified height)**.** For $1 \le i \le N-1$, define

$$\phi_i(v) := \begin{cases} \phi_{kd,\theta_{kd}}(v), & i = kd,\ 1 \le k \le d-1, \\ \|v\|^{-1}, & d \nmid i. \end{cases}$$

At the endpoints we use the convention

$$\phi_0 = \phi_N = 1.$$

For $1 \le i \le N-1$, $h \in H$ and $\Delta \in X$, set

$$\widetilde{\alpha}_{i,\eta,M}(h;\Delta) := \max\left\{ 1, \sup_{\substack{V \text{ is } \Delta\text{-rational, } \dim V = i \\ V \notin \widetilde{\mathscr{Q}}_{i,\eta,M}(\Delta) \\ 0<\|h\mathsf{w}_{\Delta,V}\|\le 1}} \phi_i(h\mathsf{w}_{\Delta,V})^{\beta_i} \right\}. \tag{6.13}$$

Finally, define

$$\widetilde{\alpha}_{\eta,M,\theta}(h;\Delta) := \bar{\alpha}_\theta(h\Delta) + \max_{1\le i\le N-1} \widetilde{\alpha}_{i,\eta,M}(h;\Delta). \tag{6.14}$$

We suppress θ from the notation when it is fixed. At degrees 0 and N, we set

$$\widehat{\alpha}_{0,\eta,M} = \widehat{\alpha}_{N,\eta,M} = \widetilde{\alpha}_{0,\eta,M} = \widetilde{\alpha}_{N,\eta,M} := 1.$$

A vector counted in (6.13) cannot lie in $\mathcal{M}_{k,1}\cup\mathcal{M}_{k,2}$: if its Plücker vector were an exact exceptional vector, then it would lie in $\mathscr{Q}_{kd,\eta,M}$, and hence the corresponding subspace would belong to $\widetilde{\mathscr{Q}}_{kd,\eta,M}(\Delta)$. Thus every value of ϕ_{kd} that occurs in (6.13) is finite.

Let

$$\beta_* := \min_{1\le i\le N-1} \beta_i = \beta_1, \qquad \tau_* := \min_{1\le i\le N-1} \tau_i > \frac{1}{2}.$$

Since the projections are fixed, $\|v - \pi_{k,r}(v)\| \ll_d \|v\|$; hence $\phi_i(v) \gg_d \|v\|^{-1}$. Terms with $\|hv\| > 1$ contribute at most 1 to the ordinary inverse height, while terms with $\|hv\| \le 1$ are included in the modified height. The definitions therefore give

$$\widehat{\alpha}_{i,\eta,M}(h;\Delta)^{\beta_i} \ll_d \widetilde{\alpha}_{i,\eta,M}(h;\Delta), \qquad \widehat{\alpha}_{\eta,M}(h;\Delta)^{\beta_*} \ll_d \widetilde{\alpha}_{\eta,M}(h;\Delta). \tag{6.15}$$

Lemma 6.2. *If $0 < \theta' \le \beta_* - 1$, then*

$$\widehat{\alpha}_{\eta,M}(h;\Delta)^{1+\theta'} \ll_d \widetilde{\alpha}_{\eta,M,\theta}(h;\Delta) \qquad ((h,\Delta) \in H \times X).$$

Proof. Since

$$\widehat{\alpha}_{\eta,M}(h;\Delta) = \max_{1\le i\le N-1} \widehat{\alpha}_{i,\eta,M}(h;\Delta),$$

and $\widehat{\alpha}_{i,\eta,M}(h;\Delta) \ge 1$, the assumption $1+\theta' \le \beta_* \le \beta_i$ gives

$$\widehat{\alpha}_{\eta,M}(h;\Delta)^{1+\theta'} \le \max_{1\le i\le N-1} \widehat{\alpha}_{i,\eta,M}(h;\Delta)^{\beta_i}.$$

The claim now follows from (6.15) and the definition of $\widetilde{\alpha}_{\eta,M,\theta}$. □

6.2. **Concavity of the exponents.** The exponents were chosen so that the products arising from intersections and sums of rational subspaces can be absorbed by the global height. The asymmetry in the first inequality below is deliberate: the intersection factor will be controlled by the ordinary height, whereas the sum factor will be controlled by the modified height. The following inequalities will be used repeatedly.

Lemma 6.3. *Let $1 \le i \le N-1$, and let $1 \le j \le \min\{i, N-i\}$.*

(1) If $i-j \ge 1$ and $i+j \le N-1$, then

$$2\beta_i^{-1} - \beta_{i+j}^{-1} - \beta_{i-j} \ge \frac{\theta_1}{2d^2}. \tag{6.16}$$

Consequently, for every $0 < \rho \le 1$ and all $x, y, z \ge 0$,

$$x^{\beta_i/2} y^{\beta_i/2} z \ll \rho\left(x^{\beta_{i+j}} + y^{1/\beta_{i-j}}\right) + 1 + \left(\rho^{-1} z\right)^{C_d\theta_1^{-1}}. \tag{6.17}$$

(2) If $i-j \ge 1$ and $i+j \le N-1$, then

$$2\tau_i^{-1} - \tau_{i+j}^{-1} - \tau_{i-j}^{-1} = 2\kappa j^2\theta_1. \tag{6.18}$$

Consequently, for every $0 < \rho \le 1$,

$$x^{\tau_i/2} y^{\tau_i/2} z \ll \rho\left(x^{\tau_{i+j}} + y^{\tau_{i-j}}\right) + 1 + \left(\rho^{-1} z\right)^{C_d\theta_1^{-1}}. \tag{6.19}$$

If exactly one of $i-j=0$ and $i+j=N$ holds, the analogous one-factor estimate is obtained by omitting the corresponding factor and height term. If both equalities hold, the product contains no height factor and is absorbed entirely into the final error term.

Proof. Put $a = (8d^2)^{-1}$. Since $\beta_r^{-1} = 1 - a\theta_r$ and $\beta_r \le 1 + 2a\theta_r$,

$$\begin{aligned} 2\beta_i^{-1} - \beta_{i+j}^{-1} - \beta_{i-j} &\ge a\big(\theta_{i+j} - 2\theta_i - 2\theta_{i-j}\big) \\ &= a\theta_i\big(10^j - 2 - 2\cdot 10^{-j}\big) \ge \frac{\theta_1}{2d^2}. \end{aligned}$$

For $\rho = 1$, (6.17) follows from weighted Young's inequality with exponents

$$\frac{2\beta_{i+j}}{\beta_i}, \qquad \frac{2}{\beta_i\beta_{i-j}}, \qquad \frac{2\beta_i^{-1}}{2\beta_i^{-1} - \beta_{i+j}^{-1} - \beta_{i-j}}.$$

The last exponent is at most $4d^2\theta_1^{-1}$. For general ρ, put

$$\gamma := \frac{\beta_i}{2\beta_{i+j}} + \frac{\beta_i\beta_{i-j}}{2} < 1$$

and apply the case $\rho = 1$ to

$$\rho^{1/\beta_{i+j}} x, \qquad \rho^{\beta_{i-j}} y, \qquad \rho^{-\gamma} z.$$

This gives (6.17), after enlarging C_d.

The identity (6.18) follows directly from $\tau_r^{-1} = 1 + \kappa r(N-r)\theta_1$. The same scaled form of weighted Young's inequality proves (6.19). With one endpoint, the argument uses two weights rather than three; with two endpoints, there is no height factor to estimate. □

6.3. **Estimates for ordinary global heights.** Put $L_s := de^s$. For $s \geq 0$, $n \in B_N(1)$, and $v \in \wedge^i \mathrm{M}_d(\mathbb{R})$, one has

$$L_s^{-2d^3}\|v\| \leq \|a_s n v\| \leq L_s^{2d^3}\|v\|. \tag{6.20}$$

At a critical degree $i = kd$, equivariance of the projections gives the corresponding estimate

$$L_s^{-2d^3}\phi_{kd}(v) \leq \phi_{kd}(a_s n v) \leq L_s^{2d^3}\phi_{kd}(v) \tag{6.21}$$

whenever $\phi_{kd}(v) < \infty$.

We first record the subharmonic estimate for $\bar{\alpha}$.

Lemma 6.4 (Margulis inequality for $\bar{\alpha}$)**.** *There are constants $c_d, C_d > 0$ such that, for every $\Delta \in X$ and $s \geq 0$,*

$$\int_{B_N(1)} \bar{\alpha}(a_s n\Delta)\, dn \ll e^{-c_d\theta_1 s}\bar{\alpha}(\Delta) + e^{C_d\theta_1^{-1} s}. \tag{6.22}$$

Proof. We first establish a local estimate. Decompose $\wedge^i \mathrm{M}_d(\mathbb{R})$ into its generic part and, when $d|i$, its two exceptional summands, using the maximum norm. If a maximal component of v is generic, Proposition 5.7 applies. If it is exceptional, apply Proposition 5.8. The estimate (6.6) places τ_i in the contracting range of that proposition. This gives the same conclusion. Thus, uniformly in $1 \leq i \leq N-1$,

$$\int_{B_N(1)} \|a_s n v\|^{-\tau_i}\, dn \ll e^{-c_d\theta_1 s}\|v\|^{-\tau_i}. \tag{6.23}$$

Fix i, and choose an i-dimensional Δ-rational subspace V with $d_\Delta(V)^{-1} = \alpha_i(\Delta)$. Such a subspace exists because the primitive decomposable vectors in $\wedge^i\Delta$ form a discrete set. Let

$$\mathcal{P}_i(\Delta, s) := \left\{ W : \begin{array}{l} W \text{ is } \Delta\text{-rational, } \dim W = i, \\ d_\Delta(W) < L_s^{4d^3} d_\Delta(V) \end{array} \right\}.$$

This set is finite, again by discreteness of the primitive exterior lattice. If $\mathcal{P}_i(\Delta, s) = \{V\}$, then (6.20) shows that V realizes the least covolume after applying $a_s n$. Hence (6.23) yields

$$\int_{B_N(1)} \alpha_i^+(a_s n\Delta)^{\tau_i}\, dn \ll e^{-c_d\theta_1 s}\alpha_i^+(\Delta)^{\tau_i} + 1.$$

Suppose next that $\mathcal{P}_i(\Delta, s)$ contains two distinct subspaces W_1, W_2. Write

$$\dim(W_1 \cap W_2) = i - j, \qquad \dim(W_1 + W_2) = i + j.$$

The Plücker covolume inequality [13, Lemma 5.6] gives

$$d_\Delta(W_1 \cap W_2) d_\Delta(W_1 + W_2) \ll d_\Delta(W_1) d_\Delta(W_2).$$

Consequently,

$$\alpha_i^+(\Delta)^2 \ll L_s^{8d^3} \alpha_{i-j}^+(\Delta) \alpha_{i+j}^+(\Delta). \tag{6.24}$$

On the other hand, (6.20) implies

$$\alpha_i^+(a_s n \Delta)^{\tau_i} \ll L_s^{2d^3} \alpha_i^+(\Delta)^{\tau_i}.$$

Combining this with (6.24), we obtain

$$\alpha_i^+(a_s n \Delta)^{\tau_i} \ll L_s^{C_d} \alpha_{i-j}^+(\Delta)^{\tau_i/2} \alpha_{i+j}^+(\Delta)^{\tau_i/2}.$$

Apply (6.19)—or its endpoint version—with $\rho = e^{-c_d \theta_1 s}$, decreasing c_d if necessary. Since each height term produced by Young's inequality is bounded by $\bar{\alpha}(\Delta)$, this gives, pointwise in n,

$$\alpha_i^+(a_s n \Delta)^{\tau_i} \ll e^{-c_d \theta_1 s} \bar{\alpha}(\Delta) + e^{C_d \theta_1^{-1} s}.$$

Summing over i proves (6.22). The implied constant may depend on d and the fixed parameter θ. □

By a similar argument, using Proposition 5.7 and Proposition 5.8 in the range of exponents slightly larger than 1, together with subcritical estimates in all other dimensions, we prove the following bounded-expansion estimate for $\bar{\alpha}'$.

Lemma 6.5 (Bounded expansion for $\bar{\alpha}'$)**.** *There is a constant $C_d > 0$ such that, for every $\Delta \in X$ and every $s \geq 0$,*

$$\int_{B_N(1)} \bar{\alpha}'_\theta(a_s n \Delta)\, dn \ll e^{C_d \theta_1 s} \bar{\alpha}'_\theta(\Delta) + e^{C_d s}. \tag{6.25}$$

The implicit constant may depend on d, but is independent of θ.

Proof. It is enough to prove the estimate for each $\bar{\alpha}'_{\theta,i}$, and then sum over $1 \leq i \leq N-1$. Fix i. The distinguished i-dimensional height is measured with exponent $1+\chi$, while the auxiliary dimensions arising from intersections and sums are measured with exponent $\sigma = \frac{3}{4}$. For sufficiently small θ, depending only on d, we have

$$1 + \chi < 1 + \frac{1}{2d^2}, \qquad \frac{1}{2} < \sigma < 1, \qquad \sigma - \frac{1+\chi}{2} > \frac{1}{5}.$$

Thus the local estimates of Section 5 apply uniformly. In the distinguished dimension i, the exponent $1+\chi$ lies in the bounded-expansion range, and the expansion loss is at most $e^{C_d \chi s}$, hence at most $e^{C_d \theta_1 s}$. In the auxiliary dimensions the exponent σ is bounded away from 1, so the corresponding local estimates have constants of size at most $e^{C_d s}$, with no loss depending on θ^{-1} or θ_1^{-1}.

We first treat the distinguished summand. Choose an i-dimensional Δ-rational subspace V with

$$d_\Delta(V)^{-1} = \alpha_i(\Delta).$$

Let $\mathcal{P}_i(\Delta, s)$ be the finite collection used in the proof of Lemma 6.4, namely

$$\mathcal{P}_i(\Delta, s) := \left\{ W : \begin{array}{l} W \text{ is } \Delta\text{-rational, } \dim W = i, \\ d_\Delta(W) < L_s^{4d^3} d_\Delta(V) \end{array} \right\}.$$

If $\mathcal{P}_i(\Delta, s) = \{V\}$, then V is the only possible minimizer after applying $a_s n$, up to the fixed log-Lipschitz factor. Applying the local estimate in degree i with exponent $1+\chi$ gives

$$\int_{B_N(1)} \alpha_i^+(a_s n\Delta)^{1+\chi}\, dn \ll e^{C_d \theta_1 s} \alpha_i^+(\Delta)^{1+\chi} + e^{C_d s}.$$

Suppose next that $\mathcal{P}_i(\Delta, s)$ contains two distinct subspaces W_1, W_2. Write

$$\dim(W_1 \cap W_2) = i - j, \qquad \dim(W_1 + W_2) = i + j,$$

with the endpoint convention when $i - j = 0$ or $i + j = N$. The intersection-sum inequality [13, Lemma 5.6] gives

$$\alpha_i^+(\Delta)^2 \ll L_s^{8d^3} \alpha_{i-j}^+(\Delta) \alpha_{i+j}^+(\Delta).$$

Therefore

$$\alpha_i^+(\Delta)^{1+\chi} \ll L_s^{C_d} \left(\alpha_{i-j}^+(\Delta) \alpha_{i+j}^+(\Delta) \right)^{(1+\chi)/2}.$$

Since $\sigma - (1+\chi)/2$ is bounded below by an absolute positive constant, the elementary scaled inequality gives

$$L_s^{C_d} \left(\alpha_{i-j}^+(\Delta) \alpha_{i+j}^+(\Delta) \right)^{(1+\chi)/2} \ll e^{C_d \theta_1 s} \left(\alpha_{i-j}^+(\Delta) \alpha_{i+j}^+(\Delta) \right)^{\sigma} + e^{C_d s}.$$

The first term on the right is one of the auxiliary summands in $\bar{\alpha}'_{\theta,i}(\Delta)$, with the required expansion factor. Hence

$$\int_{B_N(1)} \alpha_i^+(a_s n\Delta)^{1+\chi}\, dn \ll e^{C_d \theta_1 s} \bar{\alpha}'_{\theta,i}(\Delta) + e^{C_d s}.$$

It remains to control the auxiliary summands

$$\left(\alpha_{i-j}^+(\Delta) \alpha_{i+j}^+(\Delta) \right)^{\sigma}.$$

Both exponents are $\sigma = 3/4$, hence are bounded away from the critical exponent 1. Applying the same finite-competitor argument as in Lemma 6.4, but now only with these subcritical exponents, gives

$$\int_{B_N(1)} \left(\alpha_{i-j}^+(a_s n\Delta) \alpha_{i+j}^+(a_s n\Delta) \right)^{\sigma} dn \ll e^{-c_d s} \left(\alpha_{i-j}^+(\Delta) \alpha_{i+j}^+(\Delta) \right)^{\sigma} + e^{C_d s}.$$

In particular this is bounded by

$$e^{C_d \theta_1 s} \bar{\alpha}'_{\theta,i}(\Delta) + e^{C_d s}.$$

Combining the estimates for the distinguished and auxiliary summands gives

$$\int_{B_N(1)} \bar{\alpha}'_{\theta,i}(a_s n\Delta)\, dn \ll e^{C_d\theta_1 s}\bar{\alpha}'_{\theta,i}(\Delta) + e^{C_d s}.$$

Summing over $1 \le i \le N-1$ proves (6.25). □

Proposition 6.6 (Bounded expansion for the ordinary height)**.** *There is a constant $C_d > 0$ such that, for every $0 < \theta' \le 1/(2d^2)$, every $\Delta \in X$, and every $s \ge 0$,*

$$\int_{B_N(1)} \alpha(a_s n\Delta)^{1+\theta'}\, dn \ll e^{C_d\theta' s}\alpha(\Delta)^{1+\theta'} + e^{C_d s}.$$

The implicit constant may depend on d, but is independent of θ'.

Proof. It is enough to prove the estimate for each α_i^+ and then take the maximum over $1 \le i \le N-1$. Fix i, and use the same finite-competitor decomposition as in the proof of Lemma 6.5. If the relevant collection has a single competitor, the local bounded-expansion estimate in degree i, applied with exponent $1+\theta'$, gives

$$\int_{B_N(1)} \alpha_i^+(a_s n\Delta)^{1+\theta'}\, dn \ll e^{C_d\theta' s}\alpha_i^+(\Delta)^{1+\theta'} + e^{C_d s}.$$

Suppose now that two relevant i-dimensional competitors occur. If their intersection and sum have dimensions $i-j$ and $i+j$, respectively, then the standard intersection-sum inequality [13, Lemma 5.6] and the log-Lipschitz estimate give, pointwise in n,

$$\alpha_i^+(a_s n\Delta)^{1+\theta'} \ll L_s^{C_d}\left(\alpha_{i-j}^+(\Delta)\alpha_{i+j}^+(\Delta)\right)^{(1+\theta')/2},$$

with the usual endpoint convention. Applying the scaled form of Young's inequality to $X = \alpha_{i-j}^+(\Delta)^{1+\theta'}$ and $Y = \alpha_{i+j}^+(\Delta)^{1+\theta'}$, with scaling parameter $e^{C_d\theta' s}$, yields

$$L_s^{C_d}\left(\alpha_{i-j}^+(\Delta)\alpha_{i+j}^+(\Delta)\right)^{(1+\theta')/2} \ll e^{C_d\theta' s}\left(\alpha_{i-j}^+(\Delta)^{1+\theta'} + \alpha_{i+j}^+(\Delta)^{1+\theta'}\right) + e^{C_d s}.$$

The height terms on the right are bounded by $\alpha(\Delta)^{1+\theta'}$. Summing over the finitely many degrees proves the asserted estimate. □

6.4. **A consequence of the Mother Inequality.** For a critical degree $i = kd$, define

$$\bar{\phi}_i(v) := \|v - \pi_{k,1}(v)\|^{-1/2}\|v - \pi_{k,2}(v)\|^{-1/2}. \tag{6.26}$$

Then

$$\phi_i(v) = \|v\|^{-1+4\theta_i}\bar{\phi}_i(v)^{4\theta_i}. \tag{6.27}$$

We use the Mother Inequality of Benoist–Quint [2, Proposition 3.1]. Let q_λ denote the projection onto an irreducible summand of highest weight λ. In the notation relevant here, that inequality states that

$$\|q_\lambda(u)\|\,\|q_\mu(u\wedge v\wedge w)\| \ll \max_{\nu+\rho\succeq\lambda+\mu}\|q_\nu(u\wedge v)\|\,\|q_\rho(u\wedge w)\|, \tag{6.28}$$

where $\succeq$ is the dominance order and the constant is uniform over the finitely many exterior powers under consideration. Suppose now that w_1, w_2 are primitive Plücker vectors of two Δ-rational subspaces of degree kd, and let w_- and w_+ be primitive Plücker vectors of their intersection and sum. The lattice in the intersection is primitive in each of the two original lattices. We may therefore choose monomials u, v, w, adjusting signs if necessary, such that

$$u = w_-, \qquad u \wedge v = w_1, \qquad u \wedge w = w_2,$$

and

$$u \wedge v \wedge w = m w_+$$

for an integer $m \geq 1$. Applying (6.28) and discarding the factor m on the left yields

$$\|q_\lambda(w_-)\| \, \|q_\mu(w_+)\| \ll \max_{\nu+\rho \succeq \lambda+\mu} \|q_\nu(w_1)\| \, \|q_\rho(w_2)\|. \tag{6.29}$$

If $\deg w_+ < N$, then

$$\|w_-\| D_+(w_+) \ll \max\{\|w_1\| \|w_2 - \pi_{k,1}(w_2)\|, \|w_2\| \|w_1 - \pi_{k,1}(w_1)\|\}, \tag{6.30}$$

where

$$D_+(w_+) := \begin{cases} \|w_+\|, & d \nmid \deg w_+, \\ \|w_+ - \pi_{\deg(w_+)/d,1}(w_+)\|, & d \mid \deg w_+. \end{cases}$$

At any exterior degree, an irreducible summand has zero highest weight for the second SL_d-factor only when the degree is a multiple of d and the summand is the corresponding column-exceptional summand. Hence, if $q_\mu(w_+)$ lies outside the column-exceptional summand in the output degree, the second-factor component of $\lambda + \mu$ is nonzero for every weight λ occurring in the intersection degree. Since dominance is factorwise, the pair of input weights corresponding to $(\mathcal{M}_{k,1}, \mathcal{M}_{k,1})$ cannot occur in the maximum in (6.29). Thus at least one input projection is bounded by D_1 or D_2, while the other is bounded by A_2 or A_1. Taking the maximum over the finitely many intersection and output summands gives (6.30).

The next lemma is the form needed below. The restriction $\dim(W_1 + W_2) < N$ is essential; the full-space endpoint is treated separately in the proof of Theorem 6.9.

Lemma 6.7. *Let $\Delta \in X$, and let W_1, W_2 be distinct Δ-rational subspaces of dimension $i = kd$. Put*

$$\dim(W_1 \cap W_2) = i - j, \qquad \dim(W_1 + W_2) = i + j < N.$$

Let $w_m = \mathsf{w}_{\Delta, w_m}$, $m = 1, 2$, and assume that $w_m \notin \mathcal{M}_{k,1} \cup \mathcal{M}_{k,2}$ and

$$\|\pi_{k,1}(w_m)\| = \max_{q=0,1,2} \|\pi_{k,q}(w_m)\| \qquad (m = 1, 2).$$

Suppose also that, for some $\mathsf{L} \geq 1$,

$$\mathsf{L}^{-1} \|w_2\| \leq \|w_1\| \leq \mathsf{L} \|w_2\|$$

and

$$\mathsf{L}^{-1}\bar{\phi}_i(w_2) \le \bar{\phi}_i(w_1) \le \mathsf{L}\bar{\phi}_i(w_2).$$

Then

$$\phi_i(w_1)\phi_i(w_2) \ll \mathsf{L}^{8\theta_i}\|\mathsf{w}_{\Delta,W_1\cap W_2}\|^{-1}\phi_{i+j}(\mathsf{w}_{\Delta,W_1+W_2}). \tag{6.31}$$

The same conclusion holds with $\pi_{k,1}$ replaced by $\pi_{k,2}$.

Proof. We prove the first assertion. Write

$$w_- := \mathsf{w}_{\Delta,W_1\cap W_2}, \qquad w_+ := \mathsf{w}_{\Delta,W_1+W_2}.$$

Put

$$P := \max\{\|w_1\|\|w_2 - \pi_{k,1}(w_2)\|, \|w_2\|\|w_1 - \pi_{k,1}(w_1)\|\}.$$

Since the first exceptional component is maximal, norm equivalence gives

$$\bar{\phi}_i(w_m) \asymp (\|w_m\|\|w_m - \pi_{k,1}(w_m)\|)^{-1/2}.$$

The two comparison assumptions imply

$$P \ll \mathsf{L}^2\bar{\phi}_i(w_1)^{-1}\bar{\phi}_i(w_2)^{-1}. \tag{6.32}$$

The the standard intersection-sum inequality of Eskin–Margulis–Mozes [13, Lemma 5.6] gives

$$\|w_-\|\|w_+\| \ll \|w_1\|\|w_2\|, \tag{6.33}$$

and (6.30) gives

$$\|w_-\|D_+(w_+) \ll P. \tag{6.34}$$

If $d \nmid i+j$, then $D_+(w_+) = \|w_+\|$. Using (6.32)–(6.34) in (6.27), we obtain

$$\phi_i(w_1)\phi_i(w_2) \ll \mathsf{L}^{8\theta_i}(\|w_-\|\|w_+\|)^{-1} = \mathsf{L}^{8\theta_i}\|w_-\|^{-1}\phi_{i+j}(w_+).$$

Suppose that $i+j$ is divisible by d. If $D_+(w_+) = 0$, then w_+ lies in the first exceptional summand and the right-hand side of (6.31) is infinite, so there is nothing to prove. We may therefore assume that $D_+(w_+) > 0$. From (6.33) and (6.34),

$$(\|w_1\|\|w_2\|)^{-(1-4\theta_i)}P^{-4\theta_i} \ll \|w_-\|^{-1}\|w_+\|^{-1+4\theta_i}D_+(w_+)^{-4\theta_i}.$$

Since $D_+(w_+) \ll \|w_+\|$, the definition of ϕ_{i+j} and norm equivalence give

$$\phi_{i+j}(w_+) \gg \|w_+\|^{-1}\left(\frac{D_+(w_+)}{\|w_+\|}\right)^{-2\theta_{i+j}}.$$

Moreover, $\theta_{i+j} = 10^j\theta_i \ge 2\theta_i$, and $D_+(w_+)/\|w_+\|$ is bounded above by a constant depending only on the fixed projections. Hence

$$\|w_+\|^{-1+4\theta_i}D_+(w_+)^{-4\theta_i} = \|w_+\|^{-1}\left(\frac{D_+(w_+)}{\|w_+\|}\right)^{-4\theta_i} \ll \phi_{i+j}(w_+).$$

This proves (6.31). The row-exceptional case is symmetric. □

6.5. **Main global Margulis inequality for $\widetilde{\alpha}_{\eta,M}$.** Fix $0 < \eta < 1$, $M \geq 1$, $M' \geq 1$, and $B \geq 1$. For $s \geq 0$, define

$$\varepsilon_{s,\eta,M,M'}(h;\Delta) := \begin{cases} e^{-M's}, & \alpha(h\Delta) \leq L_s^{10d^3}, \\ \alpha(h\Delta)^{-BM'}, & \alpha(h\Delta) > L_s^{10d^3}. \end{cases} \tag{6.35}$$

For $1 \leq \ell \leq d-1$, let $\mathcal{E}_{\ell,s,\eta,M,M'}$ be the set of pairs $(h,\Delta) \in H \times X$ for which there is an ℓd-dimensional Δ-rational subspace V satisfying

$$\mathsf{w}_{\Delta,V} \notin \mathscr{Q}_{\ell d,\eta,M}, \qquad 0 < \|h\mathsf{w}_{\Delta,V}\| \leq L_s^{2d^3},$$

and

$$\min_{m=1,2} \|h\mathsf{w}_{\Delta,V} - \pi_{\ell,m}(h\mathsf{w}_{\Delta,V})\| \leq \varepsilon_{s,\eta,M,M'}(h;\Delta).$$

Set

$$\mathcal{E}_{s,\eta,M,M'} := \bigcup_{\ell=1}^{d-1} \mathcal{E}_{\ell,s,\eta,M,M'}. \tag{6.36}$$

The constant B will be fixed in the avoidance argument of Section 7.

We first note that the identity eventually lies outside the exceptional set.

Lemma 6.8. *Let $\Lambda \in X$, $0 < \eta \leq 1$, and $M \geq 1$. There exists $T_0 = T_0(\eta, M, \Lambda) > 0$ such that, whenever $s \geq 0$ and $s \geq T_0$,*

$$(e,\Lambda) \notin \mathcal{E}_{s,\eta,M,10d^3M}.$$

Proof. The quantity $\alpha(\Lambda)$ is fixed. Hence, for all sufficiently large s, the first line of (6.35) applies. If (e,Λ) belonged to the exceptional set, there would be a vector $v = \mathsf{w}_{\Lambda,V} \notin \mathscr{Q}_{\ell d,\eta,M}$ with $\|v\| \leq L_s^{2d^3}$ and

$$\min_{m=1,2} \|v - \pi_{\ell,m}(v)\| \leq e^{-10d^3Ms}.$$

For sufficiently large s, the right-hand side is at most $\eta L_s^{-2d^3M} \leq \eta\|v\|^{-M}$. Thus $v \in \mathscr{Q}_{\ell d,\eta,M}$, a contradiction. $\square$

We now assemble the local contraction estimates, the intersection–sum bookkeeping, and the Mother Inequality into one of the main technical ingredients of the paper. The estimate is a Margulis-type inequality rather than a pure contraction: outside the exceptional set, the averaged modified height contracts up to a logarithmic loss and an additive exponential error. The logarithmic factor reflects the possible number of competing short rational subspaces; the additive term will be absorbed later in the iteration by the choice of step sizes.

Theorem 6.9 (Main global Margulis inequality outside the exceptional set)**.** *Fix $B \geq 1$ and $M' \geq 1$. There exists $\theta_0 = \theta_0(d, B, M') > 0$ such that the following holds whenever $0 < \theta \leq \theta_0$. Let $0 < \eta < 1$, $M \geq 1$, $s \geq 0$, and*

$(h,\Delta)\notin\mathcal{E}_{s,\eta,M,M'}$. *Then*

$$\int_{B_N(1)}\widetilde{\alpha}_{\eta,M}(a_snh;\Delta)\,dn \ll e^{-c_d\theta_1 s}\widetilde{\alpha}_{\eta,M}(h;\Delta)\log\big(e+\widehat{\alpha}_{\eta,M}(h;\Delta)\big) + e^{C_d(M'\theta+\theta_1^{-1})s}. \tag{6.37}$$

The implicit constant may depend on d, B, M', and the fixed parameter θ.

Proof. Write

$$A:=\widetilde{\alpha}_{\eta,M}(h;\Delta),\quad R:=\widehat{\alpha}_{\eta,M}(h;\Delta).$$

We may assume that

$$C_dBM'\theta\le\frac{1}{20}; \tag{6.38}$$

this is achieved by reducing θ_0.

Fix $1\le i\le N-1$, and put

$$A_i=\widetilde{\alpha}_{i,\eta,M}(h;\Delta).$$

Let $\mathcal{P}_i=\mathcal{P}_i(h,\Delta,s)$ be the finite collection of i-dimensional Δ-rational subspaces V such that

$$V\notin\widetilde{\mathscr{Q}}_{i,\eta,M}(\Delta),\qquad 0<\|h\mathsf{w}_{\Delta,V}\|\le 1,$$

and

$$\phi_i(h\mathsf{w}_{\Delta,V})^{\beta_i}>L_s^{-6d^3}A_i. \tag{6.39}$$

Finiteness follows because the corresponding primitive vectors in $\wedge^i(h\Delta)$ lie in a bounded set.

We first bound subspaces outside $\mathcal{P}_i$. Let W contribute to $\widetilde{\alpha}_{i,\eta,M}(a_snh;\Delta)$, and write $w=\mathsf{w}_{\Delta,W}$, i.e. $\widetilde{\alpha}_{i,\eta,M}(a_snh;\Delta)=\phi_i(a_snhw)$. Then we have $\|hw\|\le L_s^{2d^3}$, by (6.20). If $\|hw\|\le 1$ and $W\notin\mathcal{P}_i$, then (6.21) and (6.39) give

$$\phi_i(a_snhw)^{\beta_i}\le L_s^{-2d^3}A_i$$

after enlarging the harmless constants in the definition of $\mathcal{P}_i$. If $\|hw\|>1$ and $d\nmid i$, the same expression is $O(L_s^{C_d})$. Finally, suppose $i=kd$ and $\|hw\|>1$. Since $W\notin\widetilde{\mathscr{Q}}_{i,\eta,M}(\Delta)$, its own Plücker vector does not belong to $\mathscr{Q}_{kd,\eta,M}$. Because $(h,\Delta)\notin\mathcal{E}_{s,\eta,M,M'}$,

$$\min_{m=1,2}\|hw-\pi_{k,m}(hw)\|>\varepsilon_{s,\eta,M,M'}(h;\Delta).$$

Using (6.21) and the definition of ϕ_i, we obtain

$$\phi_i(a_snhw)^{\beta_i}\ll L_s^{C_d}\left(e^{C_dM'\theta s}+\alpha(h\Delta)^{C_dBM'\theta}\right). \tag{6.40}$$

With the convention that the maximum over the empty set is zero, it follows that

$$\widetilde{\alpha}_{i,\eta,M}(a_snh;\Delta)\ll\max_{V\in\mathcal{P}_i}\phi_i(a_snh\mathsf{w}_{\Delta,V})^{\beta_i}+L_s^{-2d^3}A_i + L_s^{C_d}\left(e^{C_dM'\theta s}+\alpha(h\Delta)^{C_dBM'\theta}\right). \tag{6.41}$$

Fix a sufficiently large constant $C_0=C_0(d)$.

Case (1). First consider the case where

$$\#\mathcal{P}_i \le 3C_0 \log(e+R). \tag{6.42}$$

For noncritical i, every summand of $\wedge^i \mathrm{M}_d(\mathbb{R})$ is generic, and Proposition 5.7 applies. For $i = kd$, use Proposition 5.10. Thus

$$\int_{B_N(1)} \phi_i(a_s n h v)^{\beta_i}\, dn \ll e^{-c_d\theta_1 s}\phi_i(hv)^{\beta_i}$$

for every vector occurring in $\mathcal{P}_i$. Integrating (6.41) and using (6.42) gives the required decaying term. For the error term, the definitions of $\widehat{\alpha}$ and $\bar{\alpha}$ give $\alpha(h\Delta) \le \bar{\alpha}(h\Delta)^{1/\tau_*} \le A^{1/\tau_*}$. Together with (6.38) and $\tau_* > 1/2$, this implies

$$\alpha(h\Delta)^{C_d BM'\theta} \ll 1 + A^{1/2}.$$

The scaled inequality $uA^{1/2} \le \rho A + C\rho^{-1}u^2$, with $\rho = e^{-c_d\theta_1 s}$, absorbs this sublinear power. Consequently,

$$\int_{B_N(1)} \widetilde{\alpha}_{i,\eta,M}(a_s n h;\Delta)\, dn \ll e^{-c_d\theta_1 s} A \log(e+R) + e^{C_d(M'\theta+\theta_1^{-1})s}. \tag{6.43}$$

Case (2). Now consider the case where

$$\#\mathcal{P}_i > 3C_0 \log(e+R). \tag{6.44}$$

If $d \nmid i$, put $\mathcal{R}_i^0 = \mathcal{P}_i$ and $\mathcal{R}_i^1 = \mathcal{R}_i^2 = \varnothing$. If $i = kd$, partition $\mathcal{P}_i$ into three sets. Let $\mathcal{R}_i^0$ consist of those V for which

$$\phi_i(h\mathsf{w}_{\Delta,V}) \le L_s^{2d^3}\|h\mathsf{w}_{\Delta,V}\|^{-1}.$$

Collect the remaining subspaces in $\mathcal{R}_i^1$ or $\mathcal{R}_i^2$ according as a maximal component of their Plücker vector is the $\pi_{k,1}$- or $\pi_{k,2}$-component. At least one of these three sets contains more than $C_0\log(e+R)$ elements.
Case (2-a) $\#\mathcal{R}_i^0 \ge 2$. Let $W_1 \ne W_2$ be two subspaces in $\mathcal{R}_i^0$. Put

$$\dim(W_1 \cap W_2) = i - j, \qquad \dim(W_1 + W_2) = i + j.$$

If $i + j < N$, then $W_1 \subset W_1 + W_2$ and $W_1 \notin \widetilde{\mathscr{Q}}_{i,\eta,M}(\Delta)$, so the closure property of the downward-closed families $\widetilde{\mathscr{Q}}_{r,\eta,M}(\Delta)$ implies

$$W_1 + W_2 \notin \widetilde{\mathscr{Q}}_{i+j,\eta,M}(\Delta).$$

When $i + j = N$, we use the endpoint convention introduced above. The Plücker covolume inequality and (6.39) give

$$A_i^{2/\beta_i} \ll L_s^{C_d}\alpha_{i-j}^{+}(h\Delta)\widehat{\alpha}_{i+j,\eta,M}(h;\Delta). \tag{6.45}$$

After taking the $\beta_i/2$-power of (6.45), multiply by any fixed power of L_s needed below and absorb that power into the variable z in (6.17). Applying the scaled inequality with $\rho = e^{-c_d\theta_1 s}$ gives

$$L_s^{C_d} A_i \ll e^{-c_d\theta_1 s}\left(\widehat{\alpha}_{i+j,\eta,M}(h;\Delta)^{\beta_{i+j}} + \alpha_{i-j}^{+}(h\Delta)^{1/\beta_{i-j}}\right) + e^{C_d\theta_1^{-1}s},$$

with the absent endpoint term omitted. By (6.7) and (6.15), both displayed height terms are $\ll A$. Consequently,

$$L_s^{C_d} A_i \ll e^{-c_d\theta_1 s} A + e^{C_d\theta_1^{-1} s}. \tag{6.46}$$

This settles the case (2-a).

Case (2-b) $\#\mathcal{R}_i^0 \le 1$.

Since $\mathcal{P}_i$ is large, either $\mathcal{R}_i^1$ or $\mathcal{R}_i^2$ then contains more than $C_0 \log(e+R)$ elements. By symmetry, assume that $\mathcal{R}_i^1$ is large. Every $W \in \mathcal{R}_i^1$ satisfies

$$R^{-1} \le \|h\mathsf{w}_{\Delta,W}\| \le 1.$$

The interval $[R^{-1}, 1]$ is covered by at most $\ll \log(e+R)$ dyadic subintervals. Since C_0 was chosen sufficiently large, the pigeonhole principle gives distinct $W_1, W_2 \in \mathcal{R}_i^1$ whose Plücker norms are comparable by an absolute factor. By (6.39) and the definition of A_i, their ϕ_i-values are also comparable by a factor $L_s^{C_d}$. Using (6.27) and the comparability of their Plücker norms, we obtain

$$\mathsf{L}^{-1}\bar{\phi}_i(hw_2) \le \bar{\phi}_i(hw_1) \le \mathsf{L}\bar{\phi}_i(hw_2), \qquad \mathsf{L} \le L_s^{C_d/\theta_i}, \tag{6.47}$$

where $w_m = \mathsf{w}_{\Delta,W_m}$. Put

$$w_- := \mathsf{w}_{\Delta,W_1\cap W_2}, \qquad w_+ := \mathsf{w}_{\Delta,W_1+W_2}.$$

If $i+j<N$, then, as above, $W_1+W_2 \notin \widetilde{\mathscr{Q}}_{i+j,\eta,M}(\Delta)$.

Case (2-b-(i)) Suppose first that $i+j<N$ and $\|hw_+\| \le 1$. Apply Lemma 6.7 to the lattice $h\Delta$ and the subspaces hW_1, hW_2. Together with (6.47), this yields

$$\phi_i(hw_1)\phi_i(hw_2) \ll L_s^{C_d}\|hw_-\|^{-1}\phi_{i+j}(hw_+).$$

Combining this with (6.39) gives

$$A_i^{2/\beta_i} \ll L_s^{C_d}\alpha_{i-j}^+(h\Delta)\phi_{i+j}(hw_+).$$

Because $W_1+W_2 \notin \widetilde{\mathscr{Q}}_{i+j,\eta,M}(\Delta)$ and $\|hw_+\| \le 1$, the last factor satisfies

$$\phi_{i+j}(hw_+) \le \widetilde{\alpha}_{i+j,\eta,M}(h;\Delta)^{1/\beta_{i+j}}.$$

After multiplying by the fixed power of L_s required in (6.41), apply (6.17) with $\rho = e^{-c_d\theta_1 s}$. Using (6.7), we obtain

$$L_s^{C_d} A_i \ll e^{-c_d\theta_1 s} A + e^{C_d\theta_1^{-1} s}. \tag{6.48}$$

Case (2-b-(ii)) Now suppose that either $\|hw_+\| > 1$ or $i+j=N$. In the latter case, hw_+ is a top-degree Plücker vector of the unimodular lattice $h\Delta$, and hence $\|hw_+\| = 1$. Thus, in either case, the standard intersection-sum inequality and the comparability of $\|hw_1\|$ and $\|hw_2\|$ imply

$$\|hw_-\| \ll \|hw_1\|^2.$$

Because (h,Δ) is outside the exceptional set,

$$\|hw_1 - \pi_{k,1}(hw_1)\| > \varepsilon_{s,\eta,M,M'}(h;\Delta).$$

Using the formula for ϕ_i in the region where the first exceptional component is maximal, we obtain

$$\phi_i(hw_1) \ll \alpha_{i-j}^{+}(h\Delta)^{1/2}\left(e^{2M'\theta_i s} + \alpha(h\Delta)^{2BM'\theta_i}\right).$$

The relevance condition for W_1, together with $\alpha(h\Delta) \leq A^{1/\tau_*}$, now gives

$$A_i \ll L_s^{C_d} e^{C_d M'\theta s} A^{q_i}, \qquad q_i \leq \frac{\beta_i}{2\tau_{i-j}} + \frac{2BM'\theta_i\beta_i}{\tau_*},$$

where the first summand is omitted when $i - j = 0$. After decreasing $\theta_0(d, B, M')$, the right-hand side is at most $2/3$, uniformly in all admissible i and j. Hence

$$A_i \ll L_s^{C_d} e^{C_d M'\theta s} A^{2/3}. \tag{6.49}$$

The elementary scaled inequality $uA^{2/3} \leq \rho A + C\rho^{-2}u^3$, with $\rho = e^{-c_d\theta_1 s}$, therefore gives the same conclusion as (6.48). The case of $\mathcal{R}_i^2$ is identical.

In the large-competitor case, the log-Lipschitz estimate gives

$$\max_{V\in\mathcal{P}_i} \phi_i(a_s n h \mathsf{w}_{\Delta,V})^{\beta_i} \ll L_s^{C_d} A_i.$$

Combining this with (6.40), (6.46), (6.48), and (6.49), and absorbing the same sublinear error as in the few-competitor case, yields (6.43). Thus this bound holds for every $1 \leq i \leq N-1$. Summing over i and adding Lemma 6.4 proves (6.37). □

7. Avoidance of near-exceptional Plücker directions

The global Margulis inequality of the previous section holds outside an exceptional set $\mathcal{E}_{s,\eta,M,M'}$. This exceptional set is not the ordinary cusp. Rather, it records the event that a rational Plücker vector which is not already quasi-null for the base lattice is moved close to one of the exceptional column- or row-isotropic summands. The purpose of this section is to show that, under the Diophantine condition, this near-exceptional event has small measure in the relevant horospherical averages.

The proof has two ingredients. First, we use a small-value estimate for a prescribed set of primitive Plücker vectors. Applied to the Plücker vectors which are not already quasi-null, it shows that either the desired measure estimate holds, or else several such vectors have a wedge which remains uniformly small throughout the averaging ball. Second, the instability of nontrivial H-representations forces this wedge to be almost entirely contained in an exterior power of an exceptional summand. The Diophantine condition then upgrades this near-containment to exact containment. In the final case this would force the original Plücker vectors themselves to be quasi-null, contradicting the way the prescribed set was chosen.

Thus the argument avoids only the near-exceptional non-isotropic directions. Exact isotropic directions are not excluded here; they are omitted from the modified height and counted separately in the singular counting theorem.

We shall use two approximation parameters. Let $0 < \eta < \eta_0 < 1$, and suppose that Δ is (η_0, M)-Diophantine. Then every kd-dimensional Δ-rational subspace that is (η, M)-quasi-null is isotropic. Indeed, let $w = \mathsf{w}_{\Delta,V}$, and choose $m \in \{1, 2\}$ such that

$$\|w - \pi_{k,m}(w)\| \leq \eta \|w\|^{-M}.$$

If the left-hand side vanishes, then V is isotropic by Lemma 3.2. Otherwise the case $r = 1$ in Definition 3.6 gives

$$\eta_0 \|w\|^{-M} \leq \|w - \pi_{k,m}(w)\| \leq \eta \|w\|^{-M},$$

which contradicts $\eta < \eta_0$. Thus, under the present hypotheses, the critical quasi-null subspaces are exactly the rational isotropic critical subspaces. Consequently, the subspaces omitted by $\widehat{\alpha}_{\eta,M}$ are precisely the rational subspaces contained in such exact isotropic critical subspaces. They remain omitted; in particular, $\widehat{\alpha}_{\eta,M}$ is not identified with the ordinary Margulis height α.

Recall the definition of the exceptional set $\mathcal{E}_{s,\eta,M,M'}$ from (6.36).

Proposition 7.1 (Avoidance estimate)**.** *There exists $B = B(d) \geq 1$ such that the following holds. Let*

$$0 < \eta < \eta_0 < 1, \qquad M > 1,$$

and let $\Delta \in X$ be an (η_0, M)-Diophantine lattice. If

$$M' \geq 10d^3 BM^2, \qquad s \geq 0, \qquad t \geq 0, \qquad t \geq 10d^3 BM's,$$

then

$$\int_{B_N(1)} \alpha(a_t n\Delta)^2 \mathbb{1}_{\mathcal{E}_{s,\eta,M,M'}}(a_t n; \Delta)\, dn \ll \exp\left(-\frac{M'}{BM} s\right) + \exp\left(-\frac{1}{BM} t\right). \tag{7.1}$$

The implicit constant may depend on $d, \Delta, \eta, \eta_0,$ and M.

The remainder of the section is devoted to the proof.

7.1. **Instability and selected-vector nondivergence.** We first record a quantitative instability statement. Only the fixed ball $B_N(1)$ will be needed below.

Lemma 7.2. *Let V be a nontrivial irreducible finite-dimensional H-module. There exist constants $c = c(V) > 0$ and $\kappa = \kappa(V) > 0$ such that*

$$\sup_{n \in B_N(1)} \|a_t n v\| \geq c e^{\kappa t} \|v\|$$

for every $0 \neq v \in V$ and every $t \geq 0$.

Proof. Let A_H be the diagonal split torus of H. We first record that $V^N \neq \{0\}$ and that a_t expands V^N exponentially. Indeed, after complexification, each irreducible constituent is a tensor product of irreducible modules for the two SL_d-factors. For one factor, write the module as $\mathsf{S}_\lambda(\mathbb{C}^d)$ after adding a determinant twist, with $\lambda_d = 0$. The determinant twist is trivial on SL_d. The U_i-fixed vectors form the top component for the maximal parabolic with unipotent radical U_i. Along the principal element $b_t = \mathrm{diag}(e^{-t}, \dots, e^{-t}, e^{(d-1)t})$, every nonzero weight on this top component is at least a positive multiple of t, unless the representation of that factor is trivial. Since V is nontrivial, there are constants $c_1, \kappa_1 > 0$, depending only on V, such that

$$\|a_t w\| \geq c_1 e^{\kappa_1 t} \|w\| \qquad (w \in V^N,\ t \geq 0). \tag{7.2}$$

Let $P_N : V \to V^N$ be the projection associated with the A_H-weight decomposition. It commutes with every a_t. Define

$$q(v) := \sup_{n \in B_N(1)} \|P_N(nv)\|.$$

This is a norm on V. If $q(v) = 0$, then the polynomial map $n \mapsto P_N(nv)$ vanishes on $B_N(1)$, hence on all of N. The span of Nv is then a nonzero finite-dimensional N-module with no nonzero N-fixed vector, which is impossible for a unipotent group. Hence $q(v) = 0$ only for $v = 0$, and compactness of the unit sphere gives $q(v) \gg_V \|v\|$.

Using the boundedness of P_N, its commutation with a_t, and (7.2), we obtain

$$\sup_{n \in B_N(1)} \|a_t n v\| \gg_V \sup_{n \in B_N(1)} \|a_t P_N(nv)\| \gg_V e^{\kappa_1 t} q(v) \gg_V e^{\kappa_1 t} \|v\|.$$

This proves the lemma. □

We next state the form of quantitative non-divergence needed here. This is in the spirit of the quantitative non-divergence theorem of Kleinbock–Margulis [23], but we need a slightly refined formulation: the short vectors are required to belong to a prescribed subset of the primitive vectors, and the bad alternative must keep track of rank. More precisely, for a rank-m wedge we need the lower bound to appear at the natural scale ρ^m, rather than as a rank-independent quantity. We use the weighted-poset theorem of Kleinbock [24, Theorem 2.1], which gives exactly this form.

Recall that a continuous function f on a ball $B \subset \mathbb{R}^\ell$ is (C, τ)-*good* if, for every ball $B_0 \subset B$ and every $u > 0$,

$$\mathrm{Leb}\{x \in B_0 : |f(x)| < u\} \leq C \left(\frac{u}{\sup_{B_0} |f|} \right)^\tau \mathrm{Leb}(B_0).$$

A map $h : B \to \mathrm{GL}_k(\mathbb{R})$ is called (C, τ)-good if $x \mapsto \|h(x)w\|$ is (C, τ)-good for every $1 \leq i \leq k$ and every $w \in \wedge^i \mathbb{R}^k$.

Proposition 7.3 (Selected-vector quantitative nondivergence)**.** *For every $\ell, k \in \mathbb{N}$, there is a constant $c = c(\ell, k) > 0$ with the following property. Let $B = B(x_0, r) \subset \mathbb{R}^\ell$ be any ball. Set $\widetilde{B} = B(x_0, 3^k r)$, and let $h : \widetilde{B} \to \mathrm{GL}_k(\mathbb{R})$ be (C, τ)-good. Let $\Delta < \mathbb{R}^k$ be a lattice and $\Upsilon \subset \Delta_{\mathrm{prim}}$ be a subset. If $0 < \varepsilon \leq \rho \leq 1$, then one of the following alternatives holds:*

(1) there are linearly independent $w_1, \dots, w_m \in \Upsilon$ *for some* $1 \le m \le k$ *such that*

$$\sup_{x\in B} \|h(x)(w_1 \wedge \cdots \wedge w_m)\| \le \rho^m;$$

(2)

$$\mathrm{Leb}\left\{x \in B : \min_{w\in\Upsilon} \|h(x)w\| \le \varepsilon\right\} \le cC\left(\frac{\varepsilon}{\rho}\right)^{\tau} \mathrm{Leb}(B).$$

Proof. Let $\mathfrak{P}_\Upsilon$ be the set of nonzero rational subspaces $W < \mathbb{R}^k$ that are spanned by a finite subset of Υ, ordered by inclusion. Its length is at most k.

For $W \in \mathfrak{P}_\Upsilon$, set $\Delta_W := \Delta \cap W$, and let $\mathsf{w}_{\Delta,W}$ be a primitive Plücker vector of Δ_W. If $m = \dim W$, define

$$\iota_\Upsilon(W) := \min\left\{[\Delta_W : \mathbb{Z}w_1 + \cdots + \mathbb{Z}w_m] : \begin{array}{l} w_1, \dots, w_m \in \Upsilon \cap W, \\ \mathrm{span}_{\mathbb{R}}\{w_1, \dots, w_m\} = W \end{array}\right\}.$$

This is a positive integer. Put

$$\psi_W(x) := \iota_\Upsilon(W)\,\|h(x)\mathsf{w}_{\Delta,W}\|, \qquad \zeta(W) := \rho^{\dim W}.$$

Equivalently,

$$\psi_W(x) = \min\left\{\|h(x)(w_1 \wedge \cdots \wedge w_m)\| : \begin{array}{l} w_1, \dots, w_m \in \Upsilon \cap W, \\ \mathrm{span}_{\mathbb{R}}\{w_1, \dots, w_m\} = W \end{array}\right\}. \tag{7.3}$$

Indeed, every displayed wedge is an integral multiple of $\mathsf{w}_{\Delta,W}$, and the multiplier is precisely the corresponding index. In particular, each ψ_W is (C, τ)-good.

Apply [24, Theorem 2.1] to the weighted poset $(\mathfrak{P}_\Upsilon, \zeta)$. Its local finiteness condition holds because $\psi_W(x) < \zeta(W) \le 1$ implies $\|h(x)\mathsf{w}_{\Delta,W}\| < 1$, and a fixed lattice has only finitely many primitive subgroups of bounded covolume. If the lower-bound hypothesis of that theorem fails, then for some $W \in \mathfrak{P}_\Upsilon$,

$$\sup_{x\in B} \psi_W(x) < \rho^{\dim W}.$$

A family attaining the minimum in the definition of $\iota_\Upsilon(W)$ then gives alternative (1).

Suppose instead that the lower-bound hypothesis holds. We may assume that $\varepsilon < \rho$, since alternative (2) is trivial when $\varepsilon = \rho$. Fix λ with $1 < \lambda < \frac{\rho}{\varepsilon}$, and apply the weighted-poset theorem with marking parameter $\lambda\varepsilon/\rho$. The complement of the marked set has measure at most $cC\left(\frac{\lambda\varepsilon}{\rho}\right)^{\tau} \mathrm{Leb}(B)$. We claim that every marked point x satisfies $\|h(x)w\| > \varepsilon$ for all $w \in \Upsilon$.

Let $W_1 \subsetneq \cdots \subsetneq W_j$ be a marking flag at x, and put $W_0 = \{0\}$. Fix $w \in \Upsilon$. If $w \in W_j$, let i be the least index such that $w \in W_i$; otherwise put $i = j+1$. In both cases set $U := W_{i-1} + \mathbb{R}w$. The subspace U belongs to $\mathfrak{P}_\Upsilon$ and is comparable with every member of the flag. Thus the marking inequalities, whether U is a member of the flag or belongs to its comparable complement, give $\psi_U(x) \ge \frac{\lambda\varepsilon}{\rho}\rho^{\dim U} = \lambda\varepsilon\rho^{\dim W_{i-1}}$. Choose vectors in $\Upsilon \cap W_{i-1}$ attaining $\iota_\Upsilon(W_{i-1})$. Appending w and using (7.3), together with submultiplicativity of the exterior norm, gives $\psi_U(x) \leq \psi_{W_{i-1}}(x)\|h(x)w\| \leq \rho^{\dim W_{i-1}}\|h(x)w\|$, where the convention $\psi_{W_0} = 1$ is used when $i = 1$. Hence $\|h(x)w\| \geq \lambda\varepsilon > \varepsilon$. Therefore the sublevel set in alternative (2) is contained in the complement of the marked set. Letting $\lambda \downarrow 1$ proves alternative (2). □

7.2. **Avoiding almost-isotropic Plücker vectors.** For $0 \leq \eta < 1$, $M \geq 1$, and $1 \leq k \leq d-1$, recall the cone of Plücker vectors which are (η, M)-close to one of the exceptional summands:

$$\mathscr{Q}_{kd,\eta,M} = \left\{0 \neq w \in \wedge^{kd} \mathrm{M}_d(\mathbb{R}) : \min_{m=1,2} \|w - \pi_{k,m}(w)\| \leq \eta\|w\|^{-M}\right\}.$$

For $1 \leq k \leq d-1$, $s \geq 0$, and $0 < \varepsilon < 1$, define

$$\Xi_k(s,\varepsilon) := \left\{0 \neq v \in \wedge^{kd} \mathrm{M}_d(\mathbb{R}) : \begin{array}{l} \|v\| \leq L_s^{2d^3}, \\ \min_{m=1,2} \|v - \pi_{k,m}(v)\| \leq \varepsilon \end{array}\right\}.$$

Let $\mathcal{K}_{\eta,M}(s,\varepsilon)$ be the set of pairs $(h,\Delta) \in H \times X$ for which there exist $1 \leq k \leq d-1$ and a kd-dimensional Δ-rational subspace V such that

$$\mathsf{w}_{\Delta,V} \notin \mathscr{Q}_{kd,\eta,M} \qquad \text{and} \qquad h\mathsf{w}_{\Delta,V} \in \Xi_k(s,\varepsilon).$$

Thus $\mathcal{K}_{\eta,M}(s,\varepsilon)$ records the event that a rational Plücker vector which is not already quasi-null for the base lattice is moved close to an exceptional summand at the scale s.

The next proposition is the quantitative avoidance estimate needed in the iteration. It says that a rational Plücker vector which is not already quasi-null for the base lattice cannot be moved close to an exceptional summand for a large set of horospherical parameters. The proof combines a small-value estimate for a prescribed family of primitive Plücker vectors with instability of nontrivial H-representations; the Diophantine condition then turns near-containment in an exceptional summand into exact containment, contradicting the choice of the prescribed family.

Proposition 7.4. *There exists $B_1 = B_1(d) > 1$ such that the following holds. Let $0 < \eta < \eta_0 < 1$, $M > 1$, and let $\Delta \in X$ be (η_0, M)-Diophantine. If*

$$s \geq 0, \qquad 0 < \varepsilon < e^{-B_1 M s}, \qquad t \geq 0, \qquad t \geq B_1 \log \varepsilon^{-1},$$

then

$$\mathrm{Leb}\,\{n \in B_N(1) : (a_t n, \Delta) \in \mathcal{K}_{\eta,M}(s,\varepsilon)\} \ll \varepsilon^{1/(B_1 M)}. \tag{7.4}$$

The implicit constant may depend on Δ, η, η_0, and M.

Proof. Write $\Delta = g_0\mathbb{Z}^N$. For $1 \leq k \leq d-1$, set

$$E_k := \wedge^{kd} \mathrm{M}_d(\mathbb{R}), \qquad D_k := \dim E_k = \binom{d^2}{kd}.$$

Let

$$\Upsilon_{k,\eta,M} := \left\{\mathsf{w}_{\Delta,V} : \begin{array}{l} V \text{ is } kd\text{-dimensional and } \Delta\text{-rational}, \\ \mathsf{w}_{\Delta,V} \notin \mathscr{Q}_{kd,\eta,M} \end{array}\right\}.$$

This is a subset of the primitive vectors of the exterior lattice $\wedge^{kd}\Delta < E_k$, after choosing one sign for each Plücker vector.

Fix k. We treat proximity to $\mathcal{M}_{k,1}$; the row case is identical. Since the projections in $E_k = \mathcal{M}_{k,0} \oplus \mathcal{M}_{k,1} \oplus \mathcal{M}_{k,2}$ are fixed, one may choose $c_0 = c_0(d) > 0$ so that the linear map $\psi = \psi_{k,1,s,\varepsilon} \in \mathrm{GL}(E_k)$, defined by

$$\psi|_{\mathcal{M}_{k,1}} = c_0 \varepsilon L_s^{-2d^3}\,\mathrm{id}, \qquad \psi|_{\mathcal{M}_{k,0}\oplus\mathcal{M}_{k,2}} = c_0\,\mathrm{id},$$

has the following property: whenever $\|y\| \le L_s^{2d^3}$ and $\|y - \pi_{k,1}(y)\| \le \varepsilon$, one has $\|\psi(y)\| \le \varepsilon$. Indeed, the fixed projections have bounded operator norms, so this follows by taking c_0 sufficiently small. Define $\psi_{k,2,s,\varepsilon}$ similarly, with $\mathcal{M}_{k,2}$ in place of $\mathcal{M}_{k,1}$. By the definition of $\mathcal{K}_{\eta,M}(s,\varepsilon)$, we have

$$\begin{aligned} &\mathrm{Leb}\,\{n \in B_N(1) : (a_t n, \Delta) \in \mathcal{K}_{\eta,M}(s,\varepsilon)\} \\ &\quad \le \sum_{k=1}^{d-1}\sum_{r=1}^{2} \mathrm{Leb}\left\{ n \in B_N(1) : \min_{w\in\Upsilon_{k,\eta,M}} \|\psi_{k,r,s,\varepsilon} a_t n w\| \le \varepsilon \right\}. \end{aligned} \tag{7.5}$$

Let

$$D := \max_{1\le k\le d-1} D_k, \qquad \delta := \frac{1}{20DM}, \qquad \rho := \varepsilon^{1-\delta}.$$

Then $0 < \varepsilon \le \rho \le 1$. For every k and r, the map $n \to \psi_{k,r,s,\varepsilon} a_t n$ is (C,τ)-good on the enlarged ball required in Proposition 7.3, with $C > 0$ and $\tau > 0$ depending only on d. This follows from the multivariable Remez inequality, since all matrix coefficients in the relevant exterior powers are polynomials of degree bounded only in terms of d.

Applying Proposition 7.3, we obtain

$$\mathrm{Leb}\left\{ n \in B_N(1) : \min_{w\in\Upsilon_{k,\eta,M}} \|\psi a_t n w\| \le \varepsilon \right\} \ll \varepsilon^{\tau\delta}, \tag{7.6}$$

unless there are linearly independent $w_1, \ldots, w_m \in \Upsilon_{k,\eta,M}$, with $1 \le m \le D_k$, such that, for $\mathbf{w} := w_1 \wedge \cdots \wedge w_m \in \wedge^m E_k$, one has

$$\sup_{n\in B_N(1)} \|\psi a_t n \mathbf{w}\| \le \rho^m. \tag{7.7}$$

Here and below the same symbols denote the induced actions on exterior powers. We show that this alternative is impossible when ε is sufficiently small. The remaining compact range of ε is absorbed into the implicit constant in (7.4).

Write $E_k = \mathcal{M}_{k,1} \oplus \mathcal{M}_{k,*}$, where $\mathcal{M}_{k,*} := \mathcal{M}_{k,0} \oplus \mathcal{M}_{k,2}$, and put $q := \dim \mathcal{M}_{k,1}$. The induced decomposition

$$\wedge^m E_k = \bigoplus_j \left(\wedge^j \mathcal{M}_{k,1}\right) \wedge \left(\wedge^{m-j}\mathcal{M}_{k,*}\right) \tag{7.8}$$

is H-invariant. The map ψ acts on the j-th summand by the scalar

$$c_0^m \mathsf{r}^j, \qquad \mathsf{r} := \varepsilon L_s^{-2d^3}.$$

The common factor c_0^m is absorbed into constants depending only on d. Let P_j denote the projection onto the j-th summand. In particular,

$$P_m = \pi_{k,1}^{(m)} := \wedge^m \pi_{k,1},$$

with the convention that $P_m = 0$ if $m > q$.

Choose, in each summand of (7.8), an H-invariant complement to its H-fixed subspace. This gives an H-equivariant decomposition

$$\wedge^m E_k = V_{\mathrm{tr}} \oplus V_{\mathrm{nt}},$$

where $V_{\mathrm{tr}} = (\wedge^m E_k)^H$ and V_{nt} is a sum of nontrivial irreducible H-modules. Let P_{tr} and P_{nt} be the corresponding H-equivariant projections. By construction, they commute with ψ and with every P_j.

Since H acts trivially on V_{tr}, (7.7) implies

$$\|P_m P_{\mathrm{tr}} \mathbf{w}\| \ll \mathsf{r}^{-m} \rho^m, \tag{7.9}$$

$$\|(I - P_m) P_{\mathrm{tr}} \mathbf{w}\| \ll \mathsf{r}^{-(m-1)} \rho^m. \tag{7.10}$$

Indeed, ψ scales the top summand by r^m, and scales every off-top summand by at least r^{m-1}.

Since $m \le D$, $\delta = (20DM)^{-1}$, and $\varepsilon < e^{-B_1 M s}$, choosing $B_1 = B_1(d)$ sufficiently large gives

$$\mathsf{r}^{-m} \rho^m \ll L_s^{2d^3 m} \varepsilon^{-m\delta} \ll \varepsilon^{-1/(4M)}, \tag{7.11}$$

$$\mathsf{r}^{-(m-1)} \rho^m \ll L_s^{2d^3 (m-1)} \varepsilon^{1-m\delta} \ll \varepsilon^{1/2}. \tag{7.12}$$

On V_{nt}, the smallest singular value of the induced action of ψ is bounded below by a constant multiple of r^m. Hence (7.7) gives

$$\sup_{n \in B_N(1)} \|a_t n P_{\mathrm{nt}} \mathbf{w}\| \ll \mathsf{r}^{-m} \rho^m \ll \varepsilon^{-1/(4M)}.$$

Applying Lemma 7.2 to the finitely many irreducible summands of V_{nt}, and increasing $B_1(d)$ if necessary, the condition $t \ge B_1 \log(1/\varepsilon)$ implies

$$\|P_{\mathrm{nt}} \mathbf{w}\| \ll \varepsilon^{1/2}. \tag{7.13}$$

Combining (7.9)–(7.13), we obtain

$$0 < \|\mathbf{w}\| \ll \varepsilon^{-1/(4M)}, \qquad \|\mathbf{w} - \pi_{k,1}^{(m)}(\mathbf{w})\| \ll \varepsilon^{1/2}. \tag{7.14}$$

There is a constant $c_\Delta > 0$ such that $\|\mathbf{w}\| \ge c_\Delta$, because $\mathbf{w}$ is a nonzero vector in one of the finitely many exterior lattices $\wedge^m(\wedge^{kd}\Delta)$. If $m > q$, then $\pi_{k,1}^{(m)} = 0$, and (7.14) contradicts $\|\mathbf{w}\| \ge c_\Delta$ for sufficiently small ε.

Suppose next that $0 < m < q$, and put

$$\mathbf{w}' := \pi_{k,1}^{(m)}(\mathbf{w}) \in \wedge^m \mathcal{M}_{k,1}.$$

For sufficiently small ε, this is a nonzero decomposable vector and $\|\mathbf{w}'\| \asymp \|\mathbf{w}\|$. No nonzero decomposable line in $\wedge^m \mathcal{M}_{k,1}$ is H-fixed: otherwise the corresponding proper nonzero m-dimensional subspace of the irreducible H_1-module $\mathcal{M}_{k,1}$ would be H_1-invariant. By compactness of $\mathrm{Gr}(m, \mathcal{M}_{k,1})$, there exists $c = c(d) > 0$ such that $\|P_{\rm nt}\mathbf{w}'\| \geq c\|\mathbf{w}'\|$ for every decomposable $\mathbf{w}' \in \wedge^m \mathcal{M}_{k,1}$. On the other hand, $P_{\rm nt}$ commutes with $\pi_{k,1}^{(m)}$, so (7.13) gives $\|P_{\rm nt}\mathbf{w}'\| \ll \varepsilon^{1/2}$. Since $\|\mathbf{w}'\| \asymp \|\mathbf{w}\| \geq c_\Delta$, this is impossible for sufficiently small ε.

It remains to consider $m = q$. If $\mathbf{w} \neq \pi_{k,1}^{(q)}(\mathbf{w})$, then the (η_0, M)-Diophantine condition gives

$$\|\mathbf{w} - \pi_{k,1}^{(q)}(\mathbf{w})\| \geq \eta_0 \|\mathbf{w}\|^{-M} \gg \eta_0 \varepsilon^{1/4},$$

using the upper bound in (7.14). This contradicts the second bound in (7.14) for sufficiently small ε. Hence $\mathbf{w} = \pi_{k,1}^{(q)}(\mathbf{w})$. Since $\mathbf{w}$ is a nonzero decomposable q-vector and $\dim \mathcal{M}_{k,1} = q$, the q-dimensional subspace $\mathrm{span}\{w_1, \ldots, w_q\}$ is exactly $\mathcal{M}_{k,1}$. Hence each w_i lies in $\mathcal{M}_{k,1}$. But then

$$\min_{m=1,2} \|w_i - \pi_{k,m}(w_i)\| = 0,$$

so $w_i \in \mathscr{Q}_{kd,\eta,M}$, contradicting $w_i \in \Upsilon_{k,\eta,M}$. □

7.3. **Proof of the avoidance estimate.**

Proof of Proposition 7.1. Let $\mathsf{K} := 10d^3$. We first obtain an upper bound for $\alpha(a_t n\Delta)$ on the exceptional set. We claim that there is $C_0 = C_0(\Delta, \eta, M) \geq 1$ such that

$$(a_t n, \Delta) \in \mathcal{E}_{s,\eta,M,M'} \quad \Longrightarrow \quad \alpha(a_t n\Delta) \leq C_0 e^{t/(B^2 M)}. \tag{7.15}$$

Indeed, suppose that the pair is exceptional and that $\alpha(a_t n\Delta) > L_s^{\mathsf{K}}$, so that the second line of (6.35) applies. There are $1 \leq k \leq d-1$, $r \in \{1,2\}$, and $v = \mathsf{w}_{\Delta,V} \notin \mathscr{Q}_{kd,\eta,M}$ such that

$$0 < \|a_t n v\| \leq L_s^{2d^3}, \qquad \|a_t n v - \pi_{k,r}(a_t n v)\| \leq \alpha(a_t n\Delta)^{-BM'}.$$

The projections are H-equivariant. Hence the log-Lipschitz estimate for exterior powers gives

$$\|v\| \ll L_t^{2d^3} L_s^{2d^3}, \qquad \|v - \pi_{k,r}(v)\| \ll L_t^{2d^3} \alpha(a_t n\Delta)^{-BM'}. \tag{7.16}$$

If $\alpha(a_t n\Delta) > C_0 e^{t/(B^2 M)}$, then, upon choosing C_0 sufficiently large,

$$L_t^{2d^3} \alpha(a_t n\Delta)^{-BM'} \leq \eta \left(L_t^{2d^3} L_s^{2d^3} \right)^{-M}.$$

To verify this inequality, take logarithms and use

$$M' \geq \mathsf{K} B M^2, \qquad t \geq \mathsf{K} B M' s.$$

The negative term $-BM' \log \alpha(a_t n\Delta)$ then dominates $2d^3(M+1)\log L_t + 2d^3 M \log L_s$; the remaining constant is absorbed by C_0. Together with (7.16), this would imply

$$\|v - \pi_{k,r}(v)\| \leq \eta \|v\|^{-M},$$

contrary to the choice of v. Finally, the hypothesis $t \geq \mathsf{K} B M' s$ and the lower bound on M' imply $L_s^{\mathsf{K}} \ll e^{t/(B^2 M)}$. This proves (7.15).

For $i \geq 0$, set

$$\mathsf{h}_i := \left(de^{s+i}\right)^{\mathsf{K}}. \tag{7.17}$$

Let I be the least nonnegative integer such that

$$\mathsf{h}_I \geq C_0 e^{t/(B^2 M)}.$$

Then (7.15) implies that every exceptional point belongs to one of the following sets:

$$E_0 := \left\{ n \in B_N(1) : \begin{array}{l} \alpha(a_t n \Delta) \leq \mathsf{h}_0, \\ (a_t n, \Delta) \in \mathcal{K}_{\eta,M}(s, e^{-M's}) \end{array} \right\},$$

$$E_i := \left\{ n \in B_N(1) : \begin{array}{l} \alpha(a_t n \Delta) \leq \mathsf{h}_i, \\ (a_t n, \Delta) \in \mathcal{K}_{\eta,M}(s, \mathsf{h}_{i-1}^{-M'}) \end{array} \right\}, \qquad 1 \leq i \leq I.$$

Moreover, for a constant $I_0 = I_0(C_0, d)$,

$$I \leq I_0 + \frac{t}{\mathsf{K} B^2 M}. \tag{7.18}$$

Here we enlarged the scale $\mathsf{h}_{i-1}^{-BM'}$ occurring in (6.35) to $\mathsf{h}_{i-1}^{-M'}$, which is legitimate because $B \geq 1$.

If $s > 0$, Proposition 7.4 applies to E_0 and gives

$$\mathrm{Leb}(E_0) \ll e^{-M's/(B_1 M)}. \tag{7.19}$$

Indeed, its two hypotheses follow from $M' \geq \mathsf{K}BM^2$ and $t \geq \mathsf{K}BM's$, once B dominates B_1. When $s = 0$, the contribution of E_0 is bounded directly by h_0^2, and below we interpret (7.19) with its right-hand side replaced by a constant.

For $i \geq 1$, put $\varepsilon_i := \mathsf{h}_{i-1}^{-M'}$, and call the index i *low* if

$$B_1 \log \frac{1}{\varepsilon_i} = B_1 M' \log \mathsf{h}_{i-1} \leq \frac{t}{2}. \tag{7.20}$$

For every low index, Proposition 7.4 applies and gives

$$\mathrm{Leb}(E_i) \ll \mathsf{h}_{i-1}^{-M'/(B_1 M)}. \tag{7.21}$$

Indeed, (7.20) supplies the required lower bound for t, while $\varepsilon_i < e^{-B_1 M s}$ follows from $M' \geq \mathsf{K}BM^2$, after increasing B in terms of B_1 and d.

If $t = 0$, then $s = 0$, and the contribution is absorbed into the implicit constant. We therefore assume $t > 0$. For the remaining indices, set

$$\varepsilon_* := e^{-t/(2B_1)}. \tag{7.22}$$

If i is not low, then $\varepsilon_i < \varepsilon_*$, and hence

$$\mathcal{K}_{\eta,M}(s, \varepsilon_i) \subset \mathcal{K}_{\eta,M}(s, \varepsilon_*).$$

Moreover, $t \geq B_1 \log \frac{1}{\varepsilon_*} = \frac{t}{2}$, and the assumptions on M', s, and t imply $\varepsilon_* < e^{-B_1 M s}$, provided $B = B(d)$ is sufficiently large. Thus Proposition 7.4 also gives

$$\mathrm{Leb}(E_i) \ll e^{-t/(2B_1^2 M)} \qquad \text{for every non-low index } i. \tag{7.23}$$

From the preceding covering,

$$\int_{B_N(1)} \alpha(a_t n\Delta)^2 \mathbb{1}_{\mathcal{E}_{s,\eta,M,M'}}(a_t n;\Delta)\, dn \leq \mathsf{h}_0^2 \,\mathrm{Leb}(E_0) + \sum_{\substack{1\leq i\leq I\\ i \text{ low}}} \mathsf{h}_i^2 \,\mathrm{Leb}(E_i) + \sum_{\substack{1\leq i\leq I\\ i \text{ not low}}} \mathsf{h}_i^2 \,\mathrm{Leb}(E_i). \tag{7.24}$$

By (7.19) and the lower bound on M',

$$\mathsf{h}_0^2 \,\mathrm{Leb}(E_0) \ll e^{-M's/(BM)}. \tag{7.25}$$

Since $\mathsf{h}_i = e^{\mathsf{K}}\mathsf{h}_{i-1}$ and $M'/(B_1 M)$ is larger than a sufficiently large constant depending only on d, (7.21) gives the convergent geometric bound

$$\sum_{\substack{1\leq i\leq I\\ i \text{ low}}} \mathsf{h}_i^2 \,\mathrm{Leb}(E_i) \ll e^{-M's/(BM)}. \tag{7.26}$$

It remains to sum over the non-low indices. By the minimal choice of I,

$$\mathsf{h}_I \leq e^{\mathsf{K}} \max\left\{\mathsf{h}_0, C_0 e^{t/(B^2 M)}\right\}.$$

The assumptions relating s, t, and M' therefore give

$$\mathsf{h}_I^2 \ll \exp\left(\frac{3t}{B^2 M}\right). \tag{7.27}$$

Choose $B = B(d)$ so large that $\frac{1}{2B_1^2} - \frac{3}{B^2} > \frac{2}{B}$. Using (7.23) and (7.27), and absorbing the factor $I+1$ into half of the remaining exponential decay, we obtain

$$\sum_{\substack{1\leq i\leq I\\ i \text{ not low}}} \mathsf{h}_i^2 \,\mathrm{Leb}(E_i) \ll e^{-t/(BM)}. \tag{7.28}$$

Combining (7.24), (7.25), (7.26), and (7.28) proves (7.1). $\square$

8. Uniform boundedness of modified height functions via iterations

We now prove uniform N-average moment bounds for the modified height $\widehat{\alpha}_{\eta,M}$ along the principal $a_t N$-orbits, by combining the global Margulis-type inequality (Theorem 6.9) with the avoidance estimate (Proposition 7.1) and then iterating. The strategy is parallel to [21], but the Margulis-type inequality available here is weaker because it contains an additional logarithmic cusp term, arising from the use of the Mother Inequality of Benoist–Quint. Consequently, the iteration must separate the usual almost-isotropic exceptional set from a very deep cusp contribution and control both parts simultaneously.

In this section, we prove the following proposition, which gives the corresponding N-average estimate. It implies Theorem 3.13 after applying it with a slightly smaller approximation parameter and then replacing N-translates with K-translates.

Proposition 8.1. *Let $0 < \eta < \eta_0 < 1$, $M > 1$, and let $\Lambda \in X$ be an (η_0, M)-Diophantine lattice. Then there exists $\theta' > 0$ such that*

$$\sup_{t\geq 0} \int_{B_N(1)} \widehat{\alpha}_{\eta,M}(a_t n;\Lambda)^{1+\theta'}\, dn < \infty. \tag{8.1}$$

The proof proceeds by decomposing t into a sequence of smaller steps and iterating the global contraction estimates, while separating the contribution into good and exceptional parts.

We begin with the following standard comparison lemma for iterated horospherical averages.

Lemma 8.2. [21, Lemma 6.1] *Let $t_1, \cdots, t_m \geq 0$ for $m \in \mathbb{N}$. Then for any non-negative measurable function $f : H \times X \to [0,\infty)$ and $\Delta \in X$, we have*

$$\int_{B_N(1/3)} f(a_{t_1+\cdots+t_m} n;\Delta) dn \leq 2 \int_{B_N(1)^m} f(a_{t_m} n_m \cdots a_{t_1} n_1;\Delta) dn_1 \cdots dn_m,$$

and

$$\int_{B_N(1)^m} f(a_{t_m} n_m \cdots a_{t_1} n_1;\Delta) dn_1 \cdots dn_m \leq \int_{B_N(2)} f(a_{t_1+\cdots+t_m} n;\Delta) dn.$$

We will also use the following technical decomposition lemma from [21].

Lemma 8.3. [21, Lemma 6.2] *Let $D > 1$, $0 < \delta < \frac{1}{1+D}$, and $T > 0$ be given. For any $t \geq T/\delta$, there exists a finite sequence $\{s_i\}_{1\leq i\leq I}$ such that*

$$t = s_1 + \cdots + s_I, \qquad s_1 = D s_2, \qquad s_i = (1+\delta) s_{i+1} \quad (2 \leq i \leq I-1),$$

and $T \leq s_I \leq 2T$.

8.1. **Proof of Proposition 8.1 and Theorem 3.13.** For bounded t, the claim is immediate from the log-Lipschitz property of $\widehat{\alpha}_{\eta,M}$. Thus it suffices to consider $t \geq 0$ with t sufficiently large.

Let $B \geq 1$ be the constant from Proposition 7.1. We fix once and for all dimensional constants c_d, C_d so that the one-step estimates below hold. We then set

$$M' := 100 d^{10} B M^2, \qquad \theta = (10BM')^{-1}, \qquad D := 20 d^3 B M'.$$

The one-step estimates below follow from Proposition 6.6 and Theorem 6.9. Put

$$K_0 := C_d(M'\theta + \theta_1^{-1})$$

and choose $\theta' > 0$ so small that

$$\theta' \leq \min\left\{ \frac{\theta_1}{200d^2}, \frac{c_d \theta_1}{10000 d^3 C_d B^2 M K_0}, \frac{M'}{C_d B M(1+D)} \right\}. \tag{8.2}$$

Note that $0 < \theta' < \theta < (2d)^{-6}$. By Lemma 6.2 and the choice $\theta' \leq \theta_1/(200d^2) < \beta_* - 1$, we have

$$\widehat{\alpha}_{\eta,M}(h;\Delta)^{1+\theta'} \ll_d \widetilde{\alpha}_{\eta,M,\theta}(h;\Delta) \qquad \text{for all } (h,\Delta) \in H \times X. \tag{8.3}$$

For $s \geq 0$ define

$$\mathcal{T}_{s,\theta} := \{(h, \Delta) \in H \times X : \log \widehat{\alpha}_{\eta,M}(h; \Delta) > e^{c_d \theta_1 s}\}.$$

By Proposition 6.6, for all sufficiently large $s \geq 0$,

$$\int_{B_N(1)} \alpha(a_s n h \Delta)^{1+\theta'} \, dn \leq e^{C_d \theta' s} \alpha(h\Delta)^{1+\theta'} + e^{C_d s}. \tag{8.4}$$

Second, by Theorem 6.9, for all sufficiently large $s \geq 0$ and all $(h, \Delta) \notin \mathcal{E}_{s,\eta,M,M'} \cup \mathcal{T}_{2s,\theta}$,

$$\int_{B_N(1)} \widetilde{\alpha}_{\eta,M}(a_s n h; \Delta) \, dn \leq e^{-c_d \theta_1 s} \widetilde{\alpha}_{\eta,M}(h; \Delta) + e^{K_0 s}. \tag{8.5}$$

Moreover, by Lemma 6.8, for all sufficiently large s,

$$(\mathrm{id}, \Lambda) \notin \mathcal{E}_{s,\eta,M,10d^3 M},$$

and hence

$$\int_{B_N(1)} \widetilde{\alpha}_{\eta,M}(a_s n; \Lambda) \, dn \leq e^{-c_d \theta_1 s} \widetilde{\alpha}_{\eta,M}(\Lambda) + e^{K_0 s}. \tag{8.6}$$

Choose $T > 0$ large enough so that (8.4), (8.5), and (8.6) hold for any $s \geq 0$ with $s \geq T$. Choose $0 < \delta \leq \frac{1}{1+D}$ so that

$$C_d B \theta' M \leq M' \delta \leq \min\left\{ \frac{c_d \theta_1}{1000 d^3 B K_0}, \frac{M'}{1+D} \right\}. \tag{8.7}$$

This is possible by (8.2), after increasing the dimensional constants if necessary: the lower endpoint is at most the first upper endpoint by the second inequality in (8.2), and it is at most $M'/(1+D)$ by the third.

Now fix a sufficiently large $t \geq 0$. By Lemma 8.3, there exist a sequence $\{s_i\}_{1 \leq i \leq I}$ such that

$$t = s_1 + \cdots + s_I, \qquad s_i \geq 0,$$

and

$$s_1 = D s_2, \qquad s_i = (1+\delta) s_{i+1} \ (2 \leq i \leq I-1), \qquad T \leq s_I \leq 2T.$$

For $1 \leq i \leq I-1$, write

$$s'_{i+1} := s_{i+1} + \cdots + s_I.$$

Then Lemma 8.3 also gives

$$\left(1 - \frac{s_I}{s_i}\right) \delta^{-1} s_i \leq s'_{i+1} \leq \delta^{-1} s_i \qquad (2 \leq i \leq I-1). \tag{8.8}$$

We introduce the iterated integral

$$\mathsf{Z}_t := \int_{B_N(1)^I} \widehat{\alpha}_{\eta,M}(a_{s_I} n_I \cdots a_{s_1} n_1; \Lambda)^{1+\theta'} \, dn_1 \cdots dn_I.$$

Since $B_N(1)$ is covered by finitely many right translates of $B_N(1/3)$, by the definition of $\widehat{\alpha}_{\eta,M}$, we have

$$\int_{B_N(1)} \widehat{\alpha}_{\eta,M}(a_t n;\Lambda)^{1+\theta'}\,dn \ll \int_{B_N(1/3)} \widehat{\alpha}_{\eta,M}(a_t n;\Lambda)^{1+\theta'}\,dn,$$

and

$$\int_{B_N(1/3)} \widehat{\alpha}_{\eta,M}(a_t n;\Lambda)^{1+\theta'}\,dn \le 2\mathsf{Z}_t. \tag{8.9}$$

Therefore it suffices to show that

$$\sup_{t\ge 0} \mathsf{Z}_t < \infty. \tag{8.10}$$

For $1 \le m \le I-1$, we consider exceptional sets

$$\widetilde{\mathcal{E}}_m := \begin{cases} \mathcal{E}_{s_{m+1},\eta,M,M'} \cup \mathcal{E}_{s'_{m+1},\eta,M,M'} & \text{if } s_{m+1} \le C_d \theta_1^{-1}\log t, \\ \mathcal{E}_{s_{m+1},\eta,M,M'} & \text{otherwise.} \end{cases}$$

Define

$$\overline{\Theta}_m := \Big\{(n_1,\dots,n_m) \in B_N(1)^m : (a_{s_m} n_m \cdots a_{s_1} n_1, \Lambda) \in \widetilde{\mathcal{E}}_m\Big\},$$

and set

$$\Theta_m := \overline{\Theta}_m \times B_N(1)^{I-m} \subset B_N(1)^I, \qquad \Theta := \bigcup_{m=1}^{I-1} \Theta_m.$$

We also inductively define

$$\overline{\Omega}'_m := \Big\{(n_1,\dots,n_m) \in B_N(1)^m : (a_{s_m} n_m \cdots a_{s_1} n_1, \Lambda) \in \mathcal{T}_{s_{m+1},\theta}\Big\},$$

$$\overline{\Omega}_m := \overline{\Omega}'_m - \bigcup_{i=1}^{m-1} \left(\overline{\Omega}'_i \times B_N(1)^{m-i}\right),$$

$$\Omega_m := \overline{\Omega}_m \times B_N(1)^{I-m} \subset B_N(1)^I, \qquad \Omega := \bigcup_{m=1}^{I-1} \Omega_m.$$

We then split

$$\mathsf{Z}_t \le \mathsf{Y}_t + \sum_{m=1}^{I-1} \mathsf{T}_{t,m} + \sum_{m=1}^{I-1} \mathsf{E}_{t,m},$$

where

$$\mathsf{Y}_t := \int_{B_N(1)^I - (\Theta\cup\Omega)} \widehat{\alpha}_{\eta,M}(a_{s_I} n_I \cdots a_{s_1} n_1;\Lambda)^{1+\theta'}\,dn_1\cdots dn_I,$$

$$\mathsf{T}_{t,m} := \int_{\Omega_m - \Theta} \widehat{\alpha}_{\eta,M}(a_{s_I} n_I \cdots a_{s_1} n_1;\Lambda)^{1+\theta'}\,dn_1\cdots dn_I,$$

and

$$\mathsf{E}_{t,m} := \int_{\Theta_m} \widehat{\alpha}_{\eta,M}(a_{s_I} n_I \cdots a_{s_1} n_1;\Lambda)^{1+\theta'}\,dn_1\cdots dn_I.$$

Estimate of the bad part. Fix $1 \le m \le I-1$. Write

$$\mathsf{E}_{t,m} = \int_{\overline{\Theta}_m} J(n_1,\dots,n_m)\,dn_1\cdots dn_m,$$

where

$$J(n_1,\dots,n_m) := \int_{B_N(1)^{I-m}} \widehat{\alpha}_{\eta,M}(a_{s_I}n_I\cdots a_{s_1}n_1;\Lambda)^{1+\theta'}\,dn_{m+1}\cdots dn_I.$$

Recall that

$$\widehat{\alpha}_{\eta,M}^{1+\theta'} \le \alpha^{1+\theta'}.$$

Since $s_i \ge T$ for all $1 \le i \le I$, we may iterate the crude estimate (8.4) over the last $I-m$ steps to obtain

$$\begin{aligned} J(n_1,\dots,n_m) &\le \int_{B_N(1)^{I-m}} \alpha(a_{s_I}n_I\cdots a_{s_1}n_1\Lambda)^{1+\theta'}\,dn_{m+1}\cdots dn_I \\ &\le \sum_{i=m+1}^{I} e^{C_d s_i} e^{C_d\theta'(s_{i+1}+\cdots+s_I)}\alpha(a_{s_m}n_m\cdots a_{s_1}n_1\Lambda)^{1+\theta'} \\ &\le (I-m)e^{C_d\theta' s'_{m+2}+C_d s_{m+1}}\alpha(a_{s_m}n_m\cdots a_{s_1}n_1\Lambda)^{1+\theta'} \\ &\ll e^{(C_d\theta'\delta^{-1}+C_d)s_{m+1}}\alpha(a_{s_m}n_m\cdots a_{s_1}n_1\Lambda)^{1+\theta'}, \end{aligned} \tag{8.11}$$

using (8.8) in the third line.

Now set

$$\varphi(h;\Delta) := \alpha(h\Delta)^2\,\mathbf{1}_{\widetilde{\mathcal{E}}_m}(h,\Delta).$$

Note that $s_1 = Ds_2 \ge Ds_{m+1}$ for all $1 \le m \le I-1$, and moreover

$$s_1 = Ds_2 \ge \frac{D}{D+\delta^{-1}}t \ge D\delta^{-1}(C_d\theta_1^{-1}\log t) \ge Ds'_{m+1}$$

for sufficiently large t, if $s_{m+1} \le C_d\theta_1^{-1}\log t$. By Proposition 7.1 and Lemma 8.2,

$$\begin{aligned} &\int_{\overline{\Theta}_m} \alpha(a_{s_m}n_m\cdots a_{s_1}n_1\Lambda)^{1+\theta'}\,dn_1\cdots dn_m \\ &\le \int_{B_N(1)^m} \varphi(a_{s_m}n_m\cdots a_{s_1}n_1;\Lambda)\,dn_1\cdots dn_m \\ &\le \int_{B_N(2)} \varphi(a_{s_1+\cdots+s_m}n;\Lambda)\,dn. \end{aligned}$$

The same avoidance bound holds with $B_N(2)$ in place of $B_N(1)$, after changing the implicit constant: cover $B_N(2)$ by finitely many translates of $B_N(1)$, and use the same log-Lipschitz and goodness estimates on each fixed bounded horospherical ball. Applying Proposition 7.1 to the one or two exceptional sets appearing in $\widetilde{\mathcal{E}}_m$, and using $s'_{m+1} \ge s_{m+1}$ when the second exceptional set is present, gives

$$\int_{\overline{\Theta}_m} \alpha(a_{s_m}n_m\cdots a_{s_1}n_1\Lambda)^{1+\theta'}\,dn_1\cdots dn_m \ll e^{-\frac{M'}{BM}s_{m+1}}.$$

Here we used that $1+\theta' \le 2$. Combining this with (8.11) and using the lower bound in (8.7), we get for $1 \le m \le I-1$,

$$\mathsf{E}_{t,m} \ll \delta^{-1} \log s_{m+1} e^{\frac{M'}{5BM} s_{m+1}} e^{-\frac{M'}{BM} s_{m+1}} \ll e^{-s_{m+1}}. \tag{8.12}$$

Summing over m,

$$\sum_{m=1}^{I-1} \mathsf{E}_{t,m} \ll \sum_{m=1}^{I-1} e^{-s_{m+1}} < \sum_{m=1}^{\infty} e^{-(1+\delta)^m T} < \sum_{m=1}^{\infty} e^{-m\delta T} < \delta^{-1} T^{-1}.$$

Choose $0 < \varepsilon \le 1/100$, and let I' be the largest integer such that

$$s_{I'} \ge \varepsilon^{-1} s_I.$$

Since I' is maximal and $s_{I'} = (1+\delta)^{I-I'} s_I$, we have

$$\varepsilon^{-1} \le (1+\delta)^{I-I'} < (1+\delta)\varepsilon^{-1}.$$

In particular,

$$I - I' \asymp \delta^{-1} \log \frac{1}{\varepsilon},$$

with absolute implied constants.

Estimate of the cusp part. Fix $1 \le m \le I-1$, and put

$$x_i = (a_{s_i} n_i \cdots a_{s_1} n_1, \Lambda), \qquad 1 \le i \le m.$$

For $(n_1, \dots, n_m) \in \overline{\Omega}_m - \Theta$, we have $x_m \in \mathcal{T}_{s_{m+1},\theta}$. Hence, by the log-Lipschitz property of $\widehat{\alpha}_{\eta,M}$,

$$2d^3 t + O(1) \ge \log \widehat{\alpha}_{\eta,M}(x_m) > \exp(c_d \theta_1 s_{m+1}). \tag{8.13}$$

In particular, for sufficiently large t, $s_{m+1} \le C_d \theta_1^{-1} \log t$. By the definition of $\widetilde{\mathcal{E}}_m$, this implies that $\mathcal{E}_{s'_{m+1},\eta,M,M'} \subset \widetilde{\mathcal{E}}_m$. Since $(n_1, \dots, n_m) \notin \Theta$, we therefore have $x_m \notin \mathcal{E}_{s'_{m+1},\eta,M,M'}$.

We next show that $x_m \notin \mathcal{T}_{2s'_{m+1},\theta}$ for sufficiently large T. If $m \ge 2$, then the definition of Ω_m implies that the previous intermediate point has not entered its corresponding cusp set, i.e. $x_{m-1} \notin \mathcal{T}_{s_m,\theta}$. Thus

$$\log \widehat{\alpha}_{\eta,M}(x_{m-1}) \le \exp(c_d \theta_1 s_m).$$

Using the log-Lipschitz property of $\widehat{\alpha}_{\eta,M}$, we obtain

$$\log \widehat{\alpha}_{\eta,M}(x_m) \le \exp(c_d \theta_1 s_m) + O(s_m).$$

Since $2s'_{m+1} \ge \frac{2}{1+\delta} s_m > s_m$, the right-hand side is smaller than $\exp(c_d \theta_1 2 s'_{m+1})$ for T sufficiently large. Hence $x_m \notin \mathcal{T}_{2s'_{m+1},\theta}$. The case $m = 1$ is similar, since Λ is fixed and the log-Lipschitz property gives

$$\log \widehat{\alpha}_{\eta,M}(a_{s_1} n_1; \Lambda) \ll s_1 + 1,$$

whereas $\exp(c_d \theta_1 2 s'_2)$ grows exponentially in s'_2. Thus, for T sufficiently large, the same conclusion holds. Consequently, we have

$$x_m \notin \mathcal{E}_{s'_{m+1},\eta,M,M'} \cup \mathcal{T}_{2s'_{m+1},\theta}.$$

Therefore we may apply the contraction estimate (8.5) with $s = s'_{m+1}$.

Applying (8.5) with $s = s'_{m+1}$, we get

$$\begin{aligned}\widetilde{J}(n_1, \cdots, n_m) &= \int_{B_N(1)^{I-m}} \widetilde{\alpha}_{\eta,M}\big(a_{s_I} n_I \cdots a_{s_1} n_1; \Lambda\big) dn_{m+1} \cdots dn_I \\ &\ll \int \widetilde{\alpha}_{\eta,M}\big(a_{s'_{m+1}} n a_{s_m} n_m \cdots a_{s_1} n_1; \Lambda\big) dn \\ &\le e^{-c_d \theta_1 s'_{m+1}} \widetilde{\alpha}_{\eta,M}\big(a_{s_m} n_m \cdots a_{s_1} n_1; \Lambda\big) + e^{K_0 s'_{m+1}} \\ &\le 2e^{-c_d \theta_1 s'_{m+1}} \widetilde{\alpha}_{\eta,M}\big(a_{s_m} n_m \cdots a_{s_1} n_1; \Lambda\big).\end{aligned}$$

The last inequality follows from (8.13), because $x_m \in \mathcal{T}_{s_{m+1},\theta}$ and $s'_{m+1} \ll_\delta s_{m+1}$ for sufficiently large T, so $\widetilde{\alpha}_{\eta,M}(x_m)$ dominates the error term

$$\exp\big(K_0 s'_{m+1}\big).$$

On $\Omega_m - \Theta$, for every $1 \le i < m$, the intermediate point

$$(a_{s_i} n_i \cdots a_{s_1} n_1, \Lambda)$$

is non-exceptional for the next increment s_{i+1}:

$$(a_{s_i} n_i \cdots a_{s_1} n_1, \Lambda) \notin \mathcal{E}_{s_{i+1},\eta,M,M'} \cup \mathcal{T}_{2s_{i+1},\theta}.$$

Here the exclusion from $\mathcal{E}_{s_{i+1},\eta,M,M'}$ follows from being outside Θ, while the exclusion from $\mathcal{T}_{2s_{i+1},\theta}$ follows from the definition of Ω_m, since

$$\mathcal{T}_{2s_{i+1},\theta} \subset \mathcal{T}_{s_{i+1},\theta}.$$

Thus we may apply the contraction estimate (8.5) successively to the steps $s_m, \dots, s_2$, and then use (8.6) for the first step s_1. This gives

$$\begin{aligned}&\int_{\Omega_m - \Theta} \widetilde{\alpha}_{\eta,M}(a_{s_m} n_m \cdots a_{s_1} n_1; \Lambda)\, dn_1 \cdots dn_m \\ &\quad \le e^{-c_d\theta_1(s_1+\cdots+s_m)} \widetilde{\alpha}_{\eta,M}(\Lambda) + e^{K_0 s_1 - c_d\theta_1(s_2+\cdots+s_m)} \\ &\qquad + \sum_{i=2}^{m} e^{C_d(M'\theta+\theta_1^{-1})s_i - c_d\theta_1(s_{i+1}+\cdots+s_m)}.\end{aligned} \tag{8.14}$$

Here empty sums are interpreted as 0.

Combining (8.14) with the preceding estimate

$$\widetilde{J}(n_1, \dots, n_m) \ll e^{-c_d\theta_1 s'_{m+1}} \widetilde{\alpha}_{\eta,M}(a_{s_m} n_m \cdots a_{s_1} n_1; \Lambda),$$

and using that all s_i's have the same direction, we obtain

$$\mathsf{T}_{t,m} \ll e^{-c_d\theta_1 t} \widetilde{\alpha}_{\eta,M}(\Lambda) + e^{K_0 s_1 - c_d\theta_1 s'_2} + \sum_{i=2}^{m} e^{C_d(M'\theta+\theta_1^{-1})s_i - c_d\theta_1 s'_{i+1}}. \tag{8.15}$$

Indeed,

$$(s_{i+1} + \cdots + s_m) + s'_{m+1} = s'_{i+1}.$$

Summing (8.15) over $1 \leq m \leq I-1$, and interchanging the order of summation in the last term, gives

$$\sum_{m=1}^{I-1} \mathsf{T}_{t,m} \ll I e^{-c_d\theta_1 t}\widetilde{\alpha}_{\eta,M}(\Lambda) + I e^{K_0 s_1 - c_d\theta_1 s_2'} + \sum_{i=2}^{I-1}(I-i)e^{C_d(M'\theta+\theta_1^{-1})s_i - c_d\theta_1 s_{i+1}'}.$$

The harmless polynomial factors I and $I-i$ are absorbed by the same exponential estimates used below in (8.18)–(8.20), after increasing the initial threshold T if necessary. Hence

$$\sum_{m=1}^{I-1} \mathsf{T}_{t,m} \ll 1. \tag{8.16}$$

Estimate of the good part. On $B_N(1)^I - (\Theta \cup \Omega)$, any intermediate point

$$(a_{s_m} n_m \cdots a_{s_1} n_1, \Lambda)$$

is non-exceptional for the next increment s_{m+1}:

$$(a_{s_m} n_m \cdots a_{s_1} n_1, \Lambda) \notin \mathcal{E}_{s_{m+1},\eta,M,M'}.$$

Since the path is outside Ω, for every $1 \leq m < I$, we also have $(a_{s_m} n_m \cdots a_{s_1} n_1, \Lambda) \notin \mathcal{T}_{s_{m+1},\theta}$. Thus we may apply the contraction estimate (8.5) successively to the steps $s_I, \dots, s_2$, and then use (8.6) for the first step s_1:

$$\begin{aligned}\mathsf{Y}_t &\leq \int_{B_N(1)^I-(\Theta\cup\Omega)} \widetilde{\alpha}_{\eta,M}(a_{s_I} n_I \cdots a_{s_1} n_1; \Lambda)\, dn_1 \cdots dn_I \\ &\leq e^{-c_d\theta_1 t}\widetilde{\alpha}_{\eta,M}(\Lambda) + e^{K_0 s_1 - c_d\theta_1(s_2+\cdots+s_I)} \\ &\qquad + \sum_{m=2}^{I} e^{K_0 s_m - c_d\theta_1(s_{m+1}+\cdots+s_I)} \\ &= e^{-c_d\theta_1 t}\widetilde{\alpha}_{\eta,M}(\Lambda) + e^{K_0 s_1 - c_d\theta_1 s_2'} + \sum_{m=2}^{I} e^{K_0 s_m - c_d\theta_1 s_{m+1}'},\end{aligned} \tag{8.17}$$

where by convention $s_{I+1}' = 0$. If $2 \leq m \leq I'$, then

$$s_{m+1}' = s_{m+1} + \cdots + s_I \geq \left(1 - \frac{s_I}{s_{I'}}\right)\delta^{-1} s_m \geq (1-\varepsilon)\delta^{-1} s_m.$$

Hence, using the upper bound in (8.7), we have

$$\begin{aligned}\sum_{m=2}^{I'} e^{K_0 s_m - c_d\theta_1 s'_{m+1}} &\leq \sum_{m=2}^{I'} e^{\left(K_0 - \frac{c_d(1-\varepsilon)\delta^{-1}\theta_1}{8}\right)s_m} \\ &\leq \sum_{m=2}^{I'} e^{\left(K_0\delta - \frac{c_d(1-\varepsilon)\theta_1}{8}\right)(I-m)T} \leq \sum_{m=I-I'}^{\infty} e^{-\frac{c_d m\theta_1 T}{100}} \\ &\leq (1-e^{-\frac{c_d\theta_1 T}{100}})^{-1}\varepsilon^{\frac{c_d\delta^{-1}\theta_1 T}{100}},\end{aligned} \tag{8.18}$$

where we used $s_m \geq (1+\delta)^{I-m}T \geq (I-m)\delta T$.

For $m=1$, using the upper bound in (8.7), we get

$$c_d\theta_1 s'_2 \geq \frac{c_d(1-\varepsilon)\theta_1 s_2}{8\delta} = \frac{c_d(1-\varepsilon)\theta_1 s_1}{8D\delta} > K_0 s_1,$$

and hence

$$e^{K_0 s_1 - c_d\theta_1 s'_2} \leq 1. \tag{8.19}$$

Finally,

$$\sum_{m=I'+1}^{I} e^{K_0 s_m - c_d\theta_1(s_{m+1}+\cdots+s_I)} \leq (I-I')e^{K_0 s_I} \leq \left(\delta^{-1}\log\frac{1}{\varepsilon}\right) e^{K_0 T}. \tag{8.20}$$

Combining (8.17), (8.18), (8.19), and (8.20), we conclude that Y_t is uniformly bounded for all sufficiently large t.

Combining the bounds for the good and bad parts, we conclude that

$$\mathsf{Z}_t = \mathsf{Y}_t + \sum_{m=1}^{I-1}\mathsf{T}_{t,m} + \sum_{m=1}^{I-1}\mathsf{E}_{t,m}$$

is uniformly bounded for all sufficiently large $t \geq 0$. By (8.9), this proves

$$\sup_{t\geq 0}\int_{B_N(1)} \widehat{\alpha}_{\eta,M}(a_t n;\Lambda)^{1+\theta'}\,dn < \infty.$$

This proves Proposition 8.1.

Theorem 3.13 follows by applying Proposition 8.1 with any smaller approximation parameter in place of η, and then replacing $a_t N$-translates with $a_t K$-translates. Indeed, if $0 < \eta' < \eta$, then by Equation (3.8),

$$\widehat{\alpha}_{\eta,M}(h;\Lambda) \leq \widehat{\alpha}_{\eta',M}(h;\Lambda).$$

Using the fact that N is an expanding horospherical subgroup of G for the action of a_t, the argument of [21, Section 7.2] applies verbatim. □

We shall also use the corresponding first-moment estimate for the ordinary height. The standard iteration based on Lemma 6.4, followed by the same transfer from N-averages to K-averages, gives, for every fixed $\Delta \in X$,

$$\sup_{t\geq 0}\int_K \bar{\alpha}(a_t k\Delta)\,dk < \infty. \tag{8.21}$$

This is the usual first-moment nondivergence estimate for the standard Margulis height; compare [13, Section 7].

9. Structure of isotropic subspaces

In this section we establish the structural results on isotropic subspaces that will be used in the proof of Theorem 3.15. We first record the point at which the Diophantine hypothesis identifies quasi-null subspaces with isotropic ones. The remaining arguments are lattice-theoretic.

9.1. **Reduction to isotropic subspaces.** We first isolate the rational subspaces lying in column- or row-isotropic subspaces. For $1 \le k \le d-1$, let $\mathcal{X}^{(1)}_{k,\mathrm{iso}}(\Delta)$ and $\mathcal{X}^{(2)}_{k,\mathrm{iso}}(\Delta)$ denote respectively the collections of kd-dimensional Δ-rational column- and row-isotropic subspaces. Set

$$\mathcal{X}_{k,\mathrm{iso}}(\Delta) := \mathcal{X}^{(1)}_{k,\mathrm{iso}}(\Delta) \cup \mathcal{X}^{(2)}_{k,\mathrm{iso}}(\Delta).$$

For $m = 1, 2$, let $\mathscr{Q}^{(m)}_{k,\eta,M}(\Delta)$ denote the collection of kd-dimensional Δ-rational subspaces V such that

$$\|\mathsf{w}_{\Delta,V} - \pi_{k,m}(\mathsf{w}_{\Delta,V})\| \le \eta \|\mathsf{w}_{\Delta,V}\|^{-M}.$$

Set

$$\mathscr{Q}_{k,\eta,M}(\Delta) := \mathscr{Q}^{(1)}_{k,\eta,M}(\Delta) \cup \mathscr{Q}^{(2)}_{k,\eta,M}(\Delta).$$

We also let $\widetilde{\mathcal{X}}^{(m)}_{k,\mathrm{iso}}(\Delta)$ denote the collection of all Δ-rational subspaces contained in some member of $\mathcal{X}^{(m)}_{k,\mathrm{iso}}(\Delta)$, and set

$$\widetilde{\mathcal{X}}_{k,\mathrm{iso}}(\Delta) := \widetilde{\mathcal{X}}^{(1)}_{k,\mathrm{iso}}(\Delta) \cup \widetilde{\mathcal{X}}^{(2)}_{k,\mathrm{iso}}(\Delta).$$

Similarly, let $\widetilde{\mathscr{Q}}_{k,\eta,M}(\Delta)$ be the collection of all Δ-rational subspaces contained in some member of $\mathscr{Q}_{k,\eta,M}(\Delta)$.

Finally, define the global families

$$\mathcal{X}^{(m)}_{\mathrm{iso}}(\Delta) := \bigcup_{k=1}^{d-1} \mathcal{X}^{(m)}_{k,\mathrm{iso}}(\Delta), \qquad \mathcal{X}_{\mathrm{iso}}(\Delta) := \bigcup_{k=1}^{d-1} \mathcal{X}_{k,\mathrm{iso}}(\Delta),$$

and

$$\widetilde{\mathcal{X}}^{(m)}_{\mathrm{iso}}(\Delta) := \bigcup_{k=1}^{d-1} \widetilde{\mathcal{X}}^{(m)}_{k,\mathrm{iso}}(\Delta), \qquad \widetilde{\mathcal{X}}_{\mathrm{iso}}(\Delta) := \bigcup_{k=1}^{d-1} \widetilde{\mathcal{X}}_{k,\mathrm{iso}}(\Delta).$$

Lemma 9.1. *Let $0 < \eta < \eta_0 < 1$, let $M > 1$, and suppose that Δ is (η_0, M)-Diophantine. Then*

$$\mathscr{Q}^{(m)}_{k,\eta,M}(\Delta) = \mathcal{X}^{(m)}_{k,\mathrm{iso}}(\Delta)$$

for every $1 \le k \le d-1$ and $m \in \{1, 2\}$. Consequently,

$$\mathscr{Q}_{k,\eta,M}(\Delta) = \mathcal{X}_{k,\mathrm{iso}}(\Delta) \quad \text{and} \quad \widetilde{\mathscr{Q}}_{k,\eta,M}(\Delta) = \widetilde{\mathcal{X}}_{k,\mathrm{iso}}(\Delta).$$

Proof. Every column- or row-isotropic subspace is (η, M)-quasi-null of the corresponding type, because its Plücker vector is fixed by the corresponding exceptional projection.

Conversely, fix $m \in \{1,2\}$, and let $V \in \mathscr{Q}^{(m)}_{k,\eta,M}(\Delta)$. Put

$$w := \mathsf{w}_{\Delta,V}.$$

By definition,

$$\|w - \pi_{k,m}(w)\| \le \eta \|w\|^{-M}.$$

If $w \neq \pi_{k,m}(w)$, then the one-vector case in the (η_0, M)-Diophantine condition gives

$$\|w - \pi_{k,m}(w)\| \ge \eta_0 \|w\|^{-M},$$

contradicting $\eta < \eta_0$. Hence

$$w = \pi_{k,m}(w).$$

By Lemma 3.2, the subspace V is column-isotropic when $m = 1$, and row-isotropic when $m = 2$. This proves

$$\mathscr{Q}^{(m)}_{k,\eta,M}(\Delta) = \mathcal{X}^{(m)}_{k,\mathrm{iso}}(\Delta)$$

for each $m \in \{1,2\}$.

Taking the union over $m = 1, 2$ gives

$$\mathscr{Q}_{k,\eta,M}(\Delta) = \mathcal{X}_{k,\mathrm{iso}}(\Delta).$$

Finally, the equality of the corresponding tilde families follows by taking all Δ-rational subspaces contained in members of these equal families. □

9.2. **Orthogonal-complement duality.** Let

$$\langle X, Y \rangle := \mathrm{tr}(XY^{\mathsf{T}})$$

be the Euclidean inner product on $\mathrm{M}_d(\mathbb{R})$, and extend it to exterior powers. Fix the orientation for which the matrix units form a positively oriented orthonormal basis, and let vol be the corresponding unit volume form. If $N = d^2$, the Hodge-star

$$* : \wedge^r \mathrm{M}_d(\mathbb{R}) \to \wedge^{N-r} \mathrm{M}_d(\mathbb{R})$$

is characterized by

$$\omega \wedge (*\omega') = \langle \omega, \omega' \rangle \,\mathrm{vol}.$$

It is an isometry. We do not regard the Hodge-star itself as H-equivariant: it depends on the Euclidean inner product, which is not H-invariant. Instead, the H-equivariant duality is furnished by the wedge pairing

$$\wedge^r \mathrm{M}_d(\mathbb{R}) \otimes \wedge^{N-r} \mathrm{M}_d(\mathbb{R}) \to \wedge^N \mathrm{M}_d(\mathbb{R}).$$

Indeed, $H = \mathrm{SL}_d(\mathbb{R}) \times \mathrm{SL}_d(\mathbb{R})$ acts trivially on the top exterior power, and hence

$$\wedge^{N-r} \mathrm{M}_d(\mathbb{R}) \simeq \left(\wedge^r \mathrm{M}_d(\mathbb{R})\right)^*$$

as H-representations. In particular, when $r = kd$, the exceptional summands for k and $d - k$ are contragredient to one another, with the column and row types preserved.

We shall use the following lattice form of this duality.

Lemma 9.2. *Let $\Delta < \mathrm{M}_d(\mathbb{R})$ be a lattice, and let $V < \mathrm{M}_d(\mathbb{R})$ be a Δ-rational subspace of dimension $0 < r < d^2$. Then $V^\perp$ is Δ^*-rational and, up to sign,*

$$\mathsf{w}_{\Delta^*,V^\perp} = \frac{*\mathsf{w}_{\Delta,V}}{\mathrm{covol}(\Delta)}. \tag{9.1}$$

Consequently,

$$\mathrm{covol}_{V^\perp}(\Delta^* \cap V^\perp) = \frac{\mathrm{covol}_V(\Delta \cap V)}{\mathrm{covol}(\Delta)}. \tag{9.2}$$

Moreover, for every subspace $U < \mathbb{R}^d$,

$$\mathcal{L}(U)^\perp = \mathcal{L}(U^\perp) \quad \textit{and} \quad \mathcal{R}(U)^\perp = \mathcal{R}(U^\perp).$$

Thus orthogonal complementation preserves column and row type.

Proof. The sublattice $\Delta \cap V$ is primitive in Δ. Choose a $\mathbb{Z}$-basis $v_1, \ldots, v_r$ of $\Delta \cap V$, and extend it to a $\mathbb{Z}$-basis $v_1, \ldots, v_N$ of Δ. Let $v_1^*, \ldots, v_N^*$ be the dual basis of Δ^*. Since a vector of Δ^* lies in $V^\perp$ exactly when its first r dual coordinates vanish,

$$\Delta^* \cap V^\perp = \mathbb{Z}v_{r+1}^* \oplus \cdots \oplus \mathbb{Z}v_N^*.$$

After changing the sign of the extended basis if necessary, the standard Hodge-star identity for a basis and its dual gives

$$*\big(v_1 \wedge \cdots \wedge v_r\big) = \mathrm{covol}(\Delta)\, v_{r+1}^* \wedge \cdots \wedge v_N^*.$$

This proves (9.1), up to the unavoidable sign in the choice of Plücker vectors. Taking norms and using that the Hodge-star is an isometry yields (9.2).

Under the Euclidean tensor-product identification of matrix space,

$$\mathcal{L}(U) = U \otimes (\mathbb{R}^d)^*, \qquad \mathcal{R}(U) = \mathbb{R}^d \otimes U,$$

where the second identity uses the Euclidean identification of $\mathbb{R}^d$ with its dual. The two asserted orthogonal-complement identities follow immediately. □

9.3. **Structure in codimension d.** When Δ is (η_0, M)-Diophantine and $0 < \eta < \eta_0 < 1$, Lemma 9.1 reduces the codimension-d quasi-null family to the isotropic family. The following structural statement itself requires no Diophantine hypothesis.

Proposition 9.3. *Let $\Delta < \mathrm{M}_d(\mathbb{R})$ be a lattice with at least one Δ-rational column-isotropic subspace of dimension $d(d-1)$. Then there exist subspaces $U_1, \ldots, U_s < \mathbb{R}^d$, $s \geq 1$, such that*

$$U_1 \oplus \cdots \oplus U_s \subset \mathbb{R}^d, \qquad \dim U_j \geq 2,$$

and the following properties hold: apart from at most d exceptions, every Δ-rational column-isotropic subspace V of dimension $d(d-1)$ satisfies

$$V^\perp \subset \mathcal{L}(U_j) \quad \text{for a unique } j.$$

For every j, the lattice $\Delta^* \cap \mathcal{L}(U_j)$ is column-$\mathbb{Q}$-split; that is, it is commensurable with a tensor-product lattice $\Lambda_{1,j} \otimes \Lambda_{2,j}$, where $\Lambda_{1,j}$ is a full lattice in U_j and $\Lambda_{2,j}$ is a lattice in $(\mathbb{R}^d)^*$.

If Δ is not of $\mathbb{Q}$-split type, then every U_j is proper in $\mathbb{R}^d$. The analogous assertions hold for row-isotropic subspaces.

Proof. We prove the column statement. Let

$$\mathscr{L} := \left\{V^\perp : V \in \mathcal{X}^{(1)}_{d-1,\mathrm{iso}}(\Delta)\right\}.$$

By Lemma 9.2, each member of $\mathscr{L}$ is a Δ^*-rational subspace of the form $\mathcal{L}(\mathbb{R}u)$ for some nonzero $u \in \mathbb{R}^d$. Let

$$U := \operatorname{span}\{u : \mathcal{L}(\mathbb{R}u) \in \mathscr{L}\}.$$

Choose $L_1, \ldots, L_\ell \in \mathscr{L}$, with $Z_r = \mathcal{L}(\mathbb{R}u_r)$, so that $u_1, \ldots, u_\ell$ is a basis of U. Thus $\ell = \dim U \le d$. For each r, the inverse image of $Z_r \cap \Delta^*$ under the isomorphism $(\mathbb{R}^d)^* \to Z_r$, $\lambda \mapsto u_r \otimes \lambda$, is a lattice $\Lambda_r < (\mathbb{R}^d)^*$; hence

$$Z_r \cap \Delta^* = u_r \otimes \Lambda_r.$$

The direct sum

$$\Upsilon_0 := \bigoplus_{r=1}^{\ell} u_r \otimes \Lambda_r$$

is a full lattice in $\mathcal{L}(U)$ contained in Δ^*. The space $\mathcal{L}(U)$ is Δ^*-rational, since it is the sum of finitely many Δ^*-rational subspaces. Therefore Υ_0 and $\Upsilon := \Delta^* \cap \mathcal{L}(U)$ are commensurable. In particular, there is an integer $q \ge 1$ such that

$$\Upsilon \subset q^{-1}\Upsilon_0. \tag{9.3}$$

Define an equivalence relation on $\{1, \ldots, \ell\}$ by

$$r \sim r' \quad \Longleftrightarrow \quad c\Lambda_r \text{ is commensurable with } \Lambda_{r'} \text{ for some } c \in \mathbb{R}^\times.$$

To see how this relation controls $\mathscr{L}$, take $L = \mathcal{L}(\mathbb{R}u) \in \mathscr{L}$, and write

$$u = \sum_{r=1}^{\ell} \xi_r u_r, \qquad L \cap \Delta^* = u \otimes \Lambda$$

for a lattice $\Lambda < (\mathbb{R}^d)^*$. From (9.3), comparison of the u_r-coordinates gives

$$\xi_r \Lambda \subset q^{-1}\Lambda_r \qquad (1 \le r \le \ell).$$

Whenever $\xi_r \ne 0$, both sides are full lattices in $(\mathbb{R}^d)^*$, so $a_r\Lambda$ is commensurable with Λ_r. It follows that

$$\operatorname{supp}(u) := \{r : \xi_r \ne 0\}$$

is contained in a single $\sim$-equivalence class.

Let $E_1, \dots, E_s$ be the equivalence classes having at least two elements, and set

$$U_j := \operatorname{span}\{u_r : r \in E_j\}.$$

Because the E_j are disjoint subsets of a basis index set, $U_1 \oplus \cdots \oplus U_s$ is direct. If $\#\operatorname{supp}(u) \geq 2$, then $\operatorname{supp}(u) \subset E_j$ for a unique j, and hence $L \subset \mathcal{L}(U_j)$. If $\operatorname{supp}(u) = \{r\}$, then $L = \mathcal{Z}_r$; this line is also contained in some $\mathcal{L}(U_j)$ unless $\{r\}$ is a singleton equivalence class. Thus the only genuine exceptions are among at most $\ell \leq d$ basis lines. This proves the asserted finite decomposition and its uniqueness.

It remains to verify splitness. Fix j and choose $r_j \in E_j$. For each $r \in E_j$, choose $c_r \in \mathbb{R}^\times$ such that $c_r\Lambda_r$ is commensurable with Λ_{r_j}, and put $\widetilde{u}_r := c_r^{-1}u_r$. Then $u_r \otimes \Lambda_r = \widetilde{u}_r \otimes (c_r\Lambda_r)$ is commensurable with $\widetilde{u}_r \otimes \Lambda_{r_j}$. Consequently,

$$\bigoplus_{r \in E_j} u_r \otimes \Lambda_r$$

is commensurable with

$$\Lambda_{1,j} \otimes \Lambda_{2,j}, \qquad \Lambda_{1,j} := \bigoplus_{r \in E_j} \mathbb{Z}\widetilde{u}_r, \qquad \Lambda_{2,j} := \Lambda_{r_j}.$$

The left-hand lattice is a full sublattice of $\Delta^* \cap \mathcal{L}(U_j)$; hence $\Delta^* \cap \mathcal{L}(U_j)$ is column-$\mathbb{Q}$-split.

Finally, suppose that Δ is not of $\mathbb{Q}$-split type. If $U_j = \mathbb{R}^d$, then $\mathcal{L}(U_j) = \mathrm{M}_d(\mathbb{R})$, and the preceding conclusion would make Δ^* commensurable with a full tensor-product lattice. Dualizing would give the same conclusion for Δ, contrary to the hypothesis. Thus every U_j is proper.

The row statement follows by applying the column statement to $\Delta^\intercal$ and then transposing back. ☐

9.4. **Structure in higher codimension.** Recall the definition of the isotropic non-coincidence condition from Definition 3.14.

Proposition 9.4. *Let $\Delta < \mathrm{M}_d(\mathbb{R})$ be a lattice satisfying the column isotropic noncoincidence condition. Then, for every $\lceil \frac{d}{2} \rceil \leq k \leq d-2$, there are at most d maximal proper Δ-rational column-isotropic subspaces of dimension kd. The analogous statement holds in the row case.*

Proof. We prove the column statement. Fix $\lceil d/2 \rceil \leq k \leq d-2$, and put $\ell := d-k$. Thus $2 \leq \ell \leq \lfloor \frac{d}{2} \rfloor$. Suppose, toward a contradiction, that there are $M > d$ distinct maximal proper Δ-rational column-isotropic subspaces $V_1, \dots, V_M$ of dimension kd. Put $W_i := V_i^\perp$. By Lemma 9.2, each W_i is a Δ^*-rational column-isotropic subspace of dimension ℓd, so

$$W_i = \mathcal{L}(F_i)$$

for an ℓ-dimensional subspace $F_i < \mathbb{R}^d$.

Each W_i is minimal among the nonzero Δ^*-rational column-isotropic subspaces. Indeed, if $0 \neq Z \subsetneq W_i$ were such a subspace, then $Z^\perp$ would be a proper Δ-rational column-isotropic subspace strictly containing $V_i = W_i^{\perp}$, contradicting the maximality of V_i.

Choose a maximal subfamily $Y_1, \dots, Y_r$ of $\{W_1, \dots, W_M\}$ whose sum is direct, and set

$$\mathcal{Y} := Y_1 \oplus \cdots \oplus Y_r.$$

Write $Y_j = \mathcal{L}(E_j)$, where $\dim E_j = \ell$. The sum $E_1 \oplus \cdots \oplus E_r$ is direct, and hence

$$r\ell \leq d. \tag{9.4}$$

Moreover, every W_i is contained in $\mathcal{Y}$. Indeed, $\mathcal{Y}$ is Δ^*-rational and

$$W_i \cap \mathcal{Y} = \mathcal{L}\big(F_i \cap (E_1 \oplus \cdots \oplus E_r)\big)$$

is a Δ^*-rational column-isotropic subspace of W_i. By minimality, it is either 0 or W_i. The first alternative would allow W_i to be added to the direct family, contrary to maximality.

Since $M > d \geq r$, choose a member

$$W = \mathcal{L}(F) \in \{W_1, \dots, W_M\} - \{Y_1, \dots, Y_r\}.$$

For $1 \leq j \leq r$, put

$$L_j := Y_j \cap \Delta^*, \qquad Q := L_1 \oplus \cdots \oplus L_r.$$

The lattices Q and $\mathcal{Y} \cap \Delta^*$ are commensurable. Since $W \subset \mathcal{Y}$ is Δ^*-rational, it is also Q-rational. Let $p_j : \mathcal{Y} \to Y_j$ be the coordinate projection. If $E := E_1 \oplus \cdots \oplus E_r$, then

$$p_j = P_j \otimes \mathrm{id}_{(\mathbb{R}^d)^*}$$

for the coordinate projection $P_j : E \to E_j$. Therefore $p_j(W) = \mathcal{L}(P_j F)$ is column-isotropic. The map p_j is rational with respect to Q, so $p_j(W)$ is L_j-rational and hence Δ^*-rational. By the minimality of Y_j,

$$p_j(W) = 0 \quad \text{or} \quad p_j(W) = Y_j. \tag{9.5}$$

Choose a for which $p_a(W) = Y_a$. Since W and Y_a have the same dimension, $p_a|_W : W \to Y_a$ is an isomorphism.

For $j \neq a$, define

$$\phi_j := p_j \circ (p_a|_W)^{-1} : Y_a \to Y_j;$$

this map is zero when $p_j(W) = 0$. Then

$$W = \left\{ y + \sum_{j \neq a} \phi_j(y) : y \in Y_a \right\}.$$

Since all coordinate projections act on the first tensor factor, there are linear maps $A_j : E_a \to E_j$ such that

$$\phi_j = A_j \otimes \mathrm{id}_{(\mathbb{R}^d)^*}\,. \tag{9.6}$$

Define the shear $\Psi \in \mathrm{GL}(\mathcal{Y})$ by

$$\Psi|_{Y_a} = \mathrm{id}_{Y_a} + \sum_{j \neq a} \phi_j, \qquad \Psi|_{Y_b} = \mathrm{id}_{Y_b} \quad (b \neq a).$$

Then $\Psi(Y_a) = W$.

We claim that $\Psi(Q)$ is commensurable with Q. Since W is Q-rational, $W \cap Q$ is a full lattice in W, and

$$L'_a := p_a(W \cap Q)$$

is a full sublattice of L_a. Choose $m \geq 1$ with $mL_a \subset L'_a$. For $j \neq a$,

$$\phi_j(L'_a) \subset L_j,$$

and therefore $m\phi_j(L_a) \subset L_j$. It follows that $m\Psi(Q) \subset Q$. Since both are full lattices in $\mathcal{Y}$, $\Psi(Q)$ and Q are commensurable.

Put $A := \Psi|_{Y_a} : Y_a \to W$. The graph description and the injectivity of $p_a|_W$ give

$$A(L_a) = \Psi(Q) \cap W.$$

Because $\Psi(Q)$ and Q are commensurable, the lattices $A(L_a)$ and $Q \cap W$ are commensurable. Since Q is commensurable with $\mathcal{Y} \cap \Delta^*$, we conclude that

$$A(Y_a \cap \Delta^*) \quad \text{is commensurable with} \quad W \cap \Delta^*. \tag{9.7}$$

Let $V_a := Y_a^{\perp}$ and $V := W^{\perp}$. These are distinct members of the original family. By the quotient-lattice duality

$$\Delta^*_{V_a} = Y_a \cap \Delta^*, \qquad \Delta^*_V = W \cap \Delta^*,$$

so dualizing (9.7) yields

$$A^{-*}\Delta_{V_a} \quad \text{commensurable with} \quad \Delta_V, \qquad A^{-*} := (A^*)^{-1}.$$

By (9.6), the map A has the form

$$A = B \otimes \mathrm{id}_{(\mathbb{R}^d)^*}$$

for an isomorphism $B : E_a \to F$. Hence

$$A^{-*} = B^{-*} \otimes \mathrm{id}_{(\mathbb{R}^d)^*}.$$

Extend $B^{-*} : E_a \to F$ to an element $g \in \mathrm{GL}_d(\mathbb{R})$. Left multiplication by g on $\mathrm{M}_d(\mathbb{R})$ restricts to A^{-*} on $Y_a = \mathcal{L}(E_a)$. Therefore

$$g\Delta_{V_a} \quad \text{is commensurable with} \quad \Delta_V.$$

This contradicts the column isotropic non-coincidence condition, because $V_a \neq V$ are maximal proper Δ-rational column-isotropic subspaces of dimension kd. Thus $M > d$ is impossible.

The row case follows by transposition. □

10. Counting singular matrices: proof of Theorem 3.15

In this section we prove Theorem 3.15. The argument has three ingredients. First, we establish an asymptotic formula for bounded-rank matrices in a rational lattice. Second, we combine this formula with lattice point counting in the fibers of a rational projection. Finally, we use the structure results from Section 9 to reduce the column- and row-isotropic contributions to finitely many such fibered counting problems. The only remaining issue is to show that the overlap of the column and row contributions is of lower order. The non-isotropic singular points are treated later in Section 12; the present section counts the isotropic part Λ_{iso}.

Throughout this section, all implied constants may depend on the lattices, linear maps, and norms under consideration, but not on the parameter T.

10.1. **Bounded-rank matrices in rational lattices.** We first prove the bounded-rank counting result that will be used throughout this section. The proof decomposes matrices according to their primitive left kernels. This gives both the asymptotic formula and the summable majorants needed later for weighted limits.

Theorem 10.1 (Bounded-rank asymptotic in rational lattices)**.** *Let $1 \le k \le d-1$, and let $\Delta < \mathrm{M}_{k,d}(\mathbb{Q})$ be a full-rank lattice. For $1 \le r \le k-1$, set*

$$N_{\Delta,r}(T) := \#\{A \in \Delta : \ \|A\| \le T, \ \mathrm{rank}(A) = r\}.$$

Then there is a constant $c_{\Delta,r} > 0$ such that

$$N_{\Delta,r}(T) \sim c_{\Delta,r} T^{dr} \qquad (T \to \infty).$$

Consequently,

$$\#\{A \in \Delta : \ \|A\| \le T, \ \mathrm{rank}(A) \le r\} \sim c_{\Delta,r} T^{dr}.$$

For $\Delta = \mathrm{M}_{k,d}(\mathbb{Z})$, this was proved by Katznelson [20]. The argument below applies to any full-rank rational lattice and also gives the uniform summability estimates used in the weighted form below.

Lemma 10.2 (Primitive sublattices)**.** *Let $1 \le s \le k$. For $R \ge 3$, let $\mathcal{P}_s(k; R)$ denote the set of primitive rank-s sublattices $L < \mathbb{Z}^k$ satisfying* $\mathrm{covol}(L) \le R$. *Then*

$$\#\mathcal{P}_s(k; R) \ll R^k (\log R)^s.$$

Proof. Fix $L \in \mathcal{P}_s(k; R)$, and let $\lambda_1(L) \le \cdots \le \lambda_s(L)$ be its successive minima. Choose linearly independent vectors $v_i \in L$ with $\|v_i\| = \lambda_i(L)$. By Minkowski's second theorem [7],

$$\prod_{i=1}^{s} \|v_i\| \asymp_{k,s} \mathrm{covol}(L) \le R.$$

Since $L < \mathbb{Z}^k$, every nonzero vector in L has norm bounded below by a positive constant depending only on the chosen norm. Choose integers $m_i \ge 0$ such that $2^{m_i - 1} < \|v_i\| \le 2^{m_i}$. Then $\sum_i m_i \ll \log(2 + R)$. For a fixed tuple $(m_1, \ldots, m_s)$, the number of possible ordered tuples $(v_1, \ldots, v_s)$ is

$$\ll \prod_{i=1}^{s} 2^{k m_i} = 2^{k \sum_i m_i}.$$

The rational span of $(v_1, \ldots, v_s)$ determines L, because L is primitive. Summing over the admissible dyadic tuples gives

$$\#\mathcal{P}_s(k; R) \ll R^k (\log R)^s.$$

□

Proof of Theorem 10.1. We argue by induction on k. The assertion is empty for $k = 1$. After multiplying Δ by a nonzero integer, we may assume that $\Delta < \mathrm{M}_{k,d}(\mathbb{Z})$.

Fix $1 \le r \le k-1$, and put $s := k-r$. Let $\mathcal{P}_s(k)$ denote the set of primitive rank-s sublattices of $\mathbb{Z}^k$. For $L \in \mathcal{P}_s(k)$, define

$$V_L := \{A \in \mathrm{M}_{k,d}(\mathbb{R}) : u^\intercal A = 0 \text{ for every } u \in L\}, \qquad \Delta_L := \Delta \cap V_L.$$

Then V_L has dimension dr, and Δ_L is a full-rank lattice in V_L.

For a rational matrix A of rank r, let

$$L(A) := \ker(A^\intercal) \cap \mathbb{Z}^k.$$

Then $L(A)$ is a primitive rank-s sublattice. Conversely, if $A \in V_L$ has rank r, then $L \subset L(A)$, and primitivity together with equality of ranks gives $L = L(A)$. Hence the rank-r matrices are disjointly decomposed according to their primitive left kernels:

$$N_{\Delta,r}(T) = \sum_{L \in \mathcal{P}_s(k)} N_L(T), \tag{10.1}$$

where

$$N_L(T) := \#\{A \in \Delta_L : \ \|A\| \le T, \ \operatorname{rank}(A) = r\}.$$

For fixed L, the standard lattice-point asymptotic in the dr-dimensional space V_L gives

$$\#\{A \in \Delta_L : \ \|A\| \le T\} \sim \beta_L T^{dr}, \qquad \beta_L := \frac{\mathrm{vol}_{V_L}(B_1 \cap V_L)}{\mathrm{covol}(\Delta_L)}.$$

The matrices in Δ_L of rank at most $r-1$ identify, after an integral change of basis in $\mathbb{Z}^k$, with bounded-rank matrices in a full-rank rational lattice in $\mathrm{M}_{r,d}(\mathbb{R})$. By the induction hypothesis, their number is $O_L(T^{d(r-1)})$, with the evident interpretation when $r = 1$. Therefore

$$N_L(T) \sim \beta_L T^{dr}. \tag{10.2}$$

It remains to justify summation over L. Set

$$L_{\mathbb{Z}}^{\perp} := \{x \in \mathbb{Z}^k : u^\intercal x = 0 \text{ for every } u \in L\}.$$

Then

$$\mathrm{M}_{k,d}(\mathbb{Z}) \cap V_L = \left(L_{\mathbb{Z}}^{\perp}\right)^d \quad \text{columnwise.}$$

Since L is primitive, the Hodge-star identification of the primitive Plücker vectors of L and $L_{\mathbb{Z}}^{\perp}$ gives $\mathrm{covol}(L_{\mathbb{Z}}^{\perp}) = \mathrm{covol}(L)$. Hence

$$\mathrm{covol}\big(\mathrm{M}_{k,d}(\mathbb{Z}) \cap V_L\big) = \mathrm{covol}(L)^d.$$

Because Δ_L is a sublattice of $\mathrm{M}_{k,d}(\mathbb{Z}) \cap V_L$, we have $\mathrm{covol}(\Delta_L) \ge \mathrm{covol}(L)^d$. Since the volumes $\mathrm{vol}_{V_L}(B_1 \cap V_L)$ are uniformly bounded as V_L varies in the Grassmannian,

$$\beta_L \ll \mathrm{covol}(L)^{-d}. \tag{10.3}$$

The same geometry-of-numbers estimate gives, uniformly in L and $T \geq 1$,

$$N_L(T) \ll T^{dr} \operatorname{covol}(L)^{-d}. \tag{10.4}$$

By Lemma 10.2,

$$\sum_{L \in \mathcal{P}_s(k)} \operatorname{covol}(L)^{-d} \ll \sum_{j \geq 0} 2^{-jd} \# \mathcal{P}_s(k; 2^{j+1}) \ll \sum_{j \geq 0} 2^{-j(d-k)} (1+j)^s < \infty,$$

because $k < d$. Dominated convergence applied to (10.1), using (10.2) and (10.4), gives

$$T^{-dr} N_{\Delta,r}(T) \to \sum_{L \in \mathcal{P}_s(k)} \beta_L =: c_{\Delta,r} > 0.$$

The final assertion follows because the strata of ranks $< r$ have strictly smaller powers of T by the induction hypothesis. This proves the theorem. □

For the mixed column–row estimate below, we also need a crude bound in the square case, where this argument gives a logarithmic loss.

Corollary 10.3 (Crude bounded-rank estimate)**.** *Let $1 \leq m \leq n$, let $0 \leq r \leq m-1$, and let $T \geq 3$. Then*

$$\#\{A \in \mathrm{M}_{m,n}(\mathbb{Z}) : \ \|A\| \leq T, \ \operatorname{rank}(A) = r\} \ll T^{nr} (\log T)^{m+1}.$$

If $n > m$, the logarithmic factor may be omitted.

Proof. The case $r = 0$ is immediate. For $r \geq 1$, decompose the rank-r matrices according to their primitive rank-$(m-r)$ left kernel, as in the proof of Theorem 10.1. The contribution of a fixed kernel L is

$$\ll T^{nr} \operatorname{covol}(L)^{-n}.$$

Moreover, if such a matrix has norm at most T, then $\operatorname{covol}(L) \ll T^r$: choose r linearly independent columns and use the Hodge-star relation between the saturated column lattice and its primitive orthogonal complement. Thus it remains to sum $\operatorname{covol}(L)^{-n}$ over primitive L with $\operatorname{covol}(L) \ll T^r$. The dyadic estimate in Lemma 10.2 is summable when $n > m$, and in the square case $n = m$ gives at most $O((\log T)^{m+1})$. □

10.2. **A uniform rank-extension estimate.** We next record a uniform estimate for extending a fixed rank-deficient block while keeping the resulting matrix rank-deficient. This will be used later to bound simultaneous row- and column-rank deficiencies.

Lemma 10.4 (Uniform rank-extension estimate)**.** *Let $T \geq 3$, let $A \in \mathrm{M}_{m,a}(\mathbb{Z})$ have rank $r < m$ and norm at most T, and let $n \geq 1$. Uniformly in A,*

$$\# \left\{ B \in \mathrm{M}_{m,n}(\mathbb{Z}) : \ \|B\| \leq T, \ \operatorname{rank} \begin{pmatrix} A & B \end{pmatrix} \leq m-1 \right\} \ll \begin{cases} T^{mn}, & n < m-r, \\ T^{n(m-1)} (\log T)^m, & n \geq m-r. \end{cases}$$

The analogous estimate for vertical concatenation follows by transposition.

Proof. Let $E < \mathbb{R}^m$ be the rational span of the columns of A, let $p : \mathbb{R}^m \to E^\perp$ be the orthogonal projection, and set

$$\Upsilon_E := E \cap \mathbb{Z}^m, \qquad \overline{\Upsilon}_E := p(\mathbb{Z}^m).$$

Then Υ_E has rank r, $\overline{\Upsilon}_E$ has rank $m-r$, and

$$\operatorname{covol}(\Upsilon_E)\operatorname{covol}(\overline{\Upsilon}_E) = 1. \tag{10.5}$$

Because A contains r linearly independent columns of norm $O(T)$, all successive minima of Υ_E are $O(T)$. Hence every fiber of

$$\mathrm{M}_{m,n}(\mathbb{Z}) \to \overline{\Upsilon}_E^n, \qquad B \to p(B),$$

contains at most

$$\ll \frac{T^{rn}}{\operatorname{covol}(\Upsilon_E)^n} \tag{10.6}$$

points of norm at most T.

The rank condition is equivalent to

$$\operatorname{rank}(p(B)) \le m - r - 1. \tag{10.7}$$

If $n < m-r$, this condition is automatic, and the trivial bound $O(T^{mn})$ proves the first case.

Assume now that $n \ge m-r$, and put $q := m-r$. We claim that

$$\#\left\{\overline{B} \in \overline{\Upsilon}_E^n : \ \|\overline{B}\| \ll T, \ \operatorname{rank}(\overline{B}) \le q-1\right\} \ll \operatorname{covol}(\overline{\Upsilon}_E)^{-n} T^{n(q-1)} (\log T)^m. \tag{10.8}$$

To prove this, first consider matrices of rank $q-1$. Such matrices are contained in the union over primitive lines $\mathbb{R}u$, with $u \in \overline{\Upsilon}_E^*$, of the condition $u^\intercal \overline{B} = 0$. For fixed u, the lattice $\overline{\Upsilon}_E \cap u^\perp$ has covolume

$$\|u\| \operatorname{covol}(\overline{\Upsilon}_E),$$

and hence the contribution of the line $\mathbb{R}u$ is

$$\ll \frac{T^{n(q-1)}}{\|u\|^n \operatorname{covol}(\overline{\Upsilon}_E)^n}.$$

If this contribution is nonzero, the hyperplane lattice contains $q-1$ independent vectors of norm $O(T)$, and therefore

$$\|u\| \operatorname{covol}(\overline{\Upsilon}_E) \ll T^{q-1}.$$

Moreover, $\operatorname{covol}(\overline{\Upsilon}_E)^{-1} = \operatorname{covol}(\Upsilon_E) \ll T^r$, because Υ_E contains r independent columns of A. Thus only vectors u of norm $O(T^{m-1})$ occur. Since $\overline{\Upsilon}_E^* = E^\perp \cap \mathbb{Z}^m$ is an integral lattice of rank q, the sum over primitive u converges when $n > q$ and is $O(\log T)$ when $n = q$. The lower-rank strata are bounded in the same way, and the displayed factor $(\log T)^m$ is a convenient uniform bound for all of them. This proves (10.8).

Multiplying (10.6) and (10.8), and using (10.5), gives

$$\ll T^{rn+n(q-1)}(\log T)^m = T^{n(m-1)}(\log T)^m.$$

This proves the second case. The vertical statement follows by transposition. □

10.3. **Weighted and fibered bounded-rank asymptotics.** We now record the weighted form of the bounded-rank asymptotic and then combine it with lattice-point counting in the fibers of a rational projection. This is the form needed for the isotropic families in the next subsection.

Lemma 10.5 (Weighted bounded-rank asymptotic)**.** *Let $2 \le k \le d-1$, and let $\Delta < \mathrm{M}_{k,d}(\mathbb{Q})$ be a full-rank lattice. There is a locally finite measure μ_Δ on $\mathrm{M}_{k,d}(\mathbb{R})$, supported on the rank-$(k-1)$ locus, such that*

$$T^{-d(k-1)} \sum_{\substack{A\in\Delta \\ \mathrm{rank}(A)\le k-1}} f(A/T) \to \int_{\mathrm{M}_{k,d}(\mathbb{R})} f\, d\mu_\Delta \tag{10.9}$$

for every $f \in C_c(\mathrm{M}_{k,d}(\mathbb{R}))$.

Proof. After multiplying Δ by a nonzero integer, we may assume that $\Delta < \mathrm{M}_{k,d}(\mathbb{Z})$. For a primitive rank-one lattice $L < \mathbb{Z}^k$, let V_L and Δ_L be as in the proof of Theorem 10.1. Define

$$\mu_\Delta := \sum_{L\in\mathcal{P}_1(k)} \frac{1}{\mathrm{covol}(\Delta_L)} \mathrm{vol}_{V_L} .$$

The summability estimate (10.3) shows that this is a locally finite measure.

Matrices of rank $k-1$ have a unique primitive one-dimensional left kernel. For each fixed L, the lattice Riemann sums on Δ_L converge to $\mathrm{covol}(\Delta_L)^{-1}\,\mathrm{vol}_{V_L}$, while the rank-at-most-$(k-2)$ points in Δ_L are negligible. The majorant (10.4) permits dominated convergence over L. Finally, Theorem 10.1 shows that the total number of matrices of rank at most $k-2$ is $O(T^{d(k-2)})$, so including these matrices does not change the limit. □

Lemma 10.6 (Fibered bounded-rank asymptotic)**.** *Let*

$$\mathrm{M}_d(\mathbb{R}) = W \oplus W^\perp$$

be an orthogonal decomposition, and let $\Delta < \mathrm{M}_d(\mathbb{R})$ be a lattice. Assume that $\Delta \cap W^\perp$ is a lattice in $W^\perp$ and that $\Delta_W := p_W(\Delta)$ is a lattice in W, where p_W denotes orthogonal projection.

Suppose that $\dim W = kd$ for some $1 \le k \le d-1$, and that there is a linear isomorphism

$$\iota : W \to \mathrm{M}_{k,d}(\mathbb{R})$$

such that $\iota(\Delta_W)$ is commensurable with a full-rank lattice in $\mathrm{M}_{k,d}(\mathbb{Q})$. Then, for every norm on $\mathrm{M}_d(\mathbb{R})$, there is a constant $c_{\Delta,W} > 0$ such that

$$\#\left\{v \in \Delta :\ \|v\| \le T,\ \mathrm{rank}\big(\iota(p_W(v))\big) \le k-1\right\} \sim c_{\Delta,W} T^{d(d-1)}.$$

Proof. If $k = 1$, then the rank condition is $p_W(v) = 0$, and the assertion is the usual lattice-point asymptotic for the rank-$d(d-1)$ lattice $\Delta \cap W^\perp$. We therefore assume $k \ge 2$.

Let $\mathsf{B} := \{v \in \mathrm{M}_d(\mathbb{R}) : \|v\| \le 1\}$. For $x \in W$, define

$$\mathsf{B}_x := \{z \in W^\perp : x + z \in \mathsf{B}\}, \qquad \psi(x) := \mathrm{vol}_{W^\perp}(\mathsf{B}_x).$$

The function ψ is bounded and supported on the compact convex body $p_W(\mathsf{B})$.

For each $A \in \Delta_W$, choose $z_A \in W^\perp$ such that $A + z_A \in \Delta$. Then

$$\Delta \cap p_W^{-1}(A) = A + z_A + (\Delta \cap W^\perp).$$

The Lipschitz principle, applied in the fixed space $W^\perp$, gives uniformly in A

$$\#\left((z_A + \Delta \cap W^\perp) \cap T\mathsf{B}_{A/T}\right) = \frac{T^{d^2-kd}}{\operatorname{covol}(\Delta \cap W^\perp)}\psi(A/T) + O\big(T^{d^2-kd-1} + 1\big). \tag{10.10}$$

By Theorem 10.1, the number of relevant A with $\|A\| \ll T$ and $\operatorname{rank}(\iota(A)) \leq k-1$ is $O(T^{d(k-1)})$. Hence the total contribution of the error term in (10.10) is

$$O\left((T^{d^2-kd-1} + 1)T^{d(k-1)}\right) = o(T^{d(d-1)}).$$

Transporting Lemma 10.5 through ι gives a locally finite measure μ_{Δ_W} on W for which (10.9) holds. The discontinuities of ψ are contained in $\partial p_W(\mathsf{B})$. The measure μ_{Δ_W} is a locally finite sum of Lebesgue measures on $d(k-1)$-dimensional linear subspaces through the origin. Since 0 is an interior point of $p_W(\mathsf{B})$, none of these subspaces is contained in its boundary. Thus $\partial p_W(\mathsf{B})$ is μ_{Δ_W}-null, and the weighted convergence applies to ψ by approximation from above and below.

Summing (10.10) over the bounded-rank points in Δ_W gives

$$\begin{aligned}
&\#\left\{v \in \Delta : \ \|v\| \leq T, \ \operatorname{rank}\big(\iota(p_W(v))\big) \leq k-1\right\} \\
&\quad = \frac{T^{d^2-kd}}{\operatorname{covol}(\Delta \cap W^\perp)} \sum_{\substack{A \in \Delta_W \\ \operatorname{rank}(\iota(A)) \leq k-1}} \psi(A/T) + o(T^{d(d-1)}) \\
&\quad \sim \frac{1}{\operatorname{covol}(\Delta \cap W^\perp)} \left(\int_W \psi \, d\mu_{\Delta_W}\right) T^{d(d-1)}.
\end{aligned}$$

The integral is positive because ψ is positive near the origin on each linear subspace appearing in the support of μ_{Δ_W}. This proves the lemma. □

10.4. **Finite decompositions of the isotropic contribution.** Let $\Lambda_{\mathrm{iso}}^{(1)}$ (or $\Lambda_{\mathrm{iso}}^{(2)}$) denote the set of lattice points that lie in a Λ-rational column (or row)-isotropic subspace of one of the dimensions $d, 2d, \ldots, d(d-1)$. Recall that Λ_{iso} denote the set of lattice points that lie in a Λ-rational column- or row-isotropic subspace. Thus

$$\Lambda_{\mathrm{iso}} = \Lambda_{\mathrm{iso}}^{(1)} \cup \Lambda_{\mathrm{iso}}^{(2)}. \tag{10.11}$$

The next lemma packages the structural results of Section 9 in the form needed for counting. It says that, up to lower-dimensional pieces, the column-isotropic contribution is a finite union of fibered bounded-rank conditions.

Lemma 10.7 (Finite piece decomposition)**.** *Assume that Λ is not $\mathbb{Q}$-split and satisfies the column isotropic noncoincidence condition. Then there are*

(i) *finitely many proper Λ-rational subspaces $L_1, \dots, L_q$, each contained in a column-isotropic subspace and each of dimension at most $d(d-2)$; and*
(ii) *finitely many column-type subspaces*

$$W_j = \mathcal{L}(U_j) = U_j \otimes (\mathbb{R}^d)^*, \qquad 1 \le k_j := \dim U_j \le d-1,$$

together with rank-preserving tensor-product isomorphisms

$$\iota_j : W_j \to \mathrm{M}_{k_j,d}(\mathbb{R}),$$

such that $\iota_j(p_{W_j}(\Lambda))$ is commensurable with a full-rank rational lattice,

for which, on writing

$$\Sigma_j := \left\{ v \in \Lambda : \operatorname{rank}\bigl(\iota_j(p_{W_j}(v))\bigr) \le k_j - 1 \right\},$$

one has

$$\Lambda^{(1)}_{\mathrm{iso}} = \left(\Lambda \cap \bigcup_{m=1}^{q} L_m \right) \cup \bigcup_{j=1}^{s} \Sigma_j. \tag{10.12}$$

Moreover, the pieces may be chosen so that, whenever $j \neq j'$,

$$\#\{v \in \Sigma_j \cap \Sigma_{j'} : \|v\| \le T\} \ll T^{d(d-2)}. \tag{10.13}$$

The analogous statement holds for the row-isotropic contribution.

Proof. We prove the column statement. First consider maximal proper Λ-rational column-isotropic subspaces of dimension kd with $k \le d-2$. For $\lceil d/2 \rceil \le k \le d-2$, their finiteness follows from Proposition 9.4. For $k < d/2$, there is at most one such maximal subspace for each k. Indeed, if V_1 and V_2 were distinct maximal subspaces of this dimension, then $V_1^\perp$ and $V_2^\perp$ would be minimal nonzero Λ^*-rational column-isotropic subspaces of dimension $(d-k)d$. The corresponding $(d-k)$-dimensional subspaces of $\mathbb{R}^d$ intersect nontrivially, because $d-k > d/2$. Their intersection is again Λ^*-rational and column-isotropic, contradicting minimality unless $V_1^\perp = V_2^\perp$. Taking all maximal subspaces in these finitely many families gives $L_1, \dots, L_q$, and we include the zero subspace if necessary. Every column-isotropic point not contained in a $d(d-1)$-dimensional column-isotropic subspace lies in one of these L_m.

It remains to treat the $d(d-1)$-dimensional column-isotropic subspaces. By Proposition 9.3, and using the construction in its proof, there are pairwise linearly disjoint subspaces $U_1, \dots, U_s$, of dimensions between 2 and $d-1$, such that every Λ-rational column-isotropic subspace V of dimension $d(d-1)$, apart from finitely many exceptional ones, satisfies

$$V^\perp \subset \mathcal{L}(U_j)$$

for a unique j. Moreover, $\Lambda^* \cap \mathcal{L}(U_j)$ is column-$\mathbb{Q}$-split. We discard from the exceptional list any member whose orthogonal complement is already contained in one of the $\mathcal{L}(U_j)$. For each remaining exceptional subspace V, add the one-dimensional factor U determined by $V^\perp = \mathcal{L}(U)$ to the list.

Thus every $d(d-1)$-dimensional column-isotropic subspace is assigned to one of the resulting spaces $W_j = \mathcal{L}(U_j)$, possibly with $k_j = 1$.

Fix one such space $W_j = \mathcal{L}(U_j)$, and write $p_j = p_{W_j}$. Since

$$p_j(\Lambda)^* = \Lambda^* \cap W_j,$$

and tensor-product splitness is preserved by duality, there is a rank-preserving tensor-product isomorphism

$$\iota_j = a_j \otimes b_j : W_j \to \mathrm{M}_{k_j,d}(\mathbb{R})$$

for which $\iota_j(p_j(\Lambda))$ is commensurable with a full-rank rational lattice. When $k_j = 1$, this is automatic.

We claim that

$$\Lambda \cap \bigcup_{\substack{V \in \mathcal{X}^{(1)}_{d-1,\mathrm{iso}}(\Lambda) \\ V^\perp \subset W_j}} V = \Sigma_j. \tag{10.14}$$

If $v \in V \cap \Lambda$ and $V^\perp = \mathcal{L}(\mathbb{R}u)$ with $0 \neq u \in U_j$, then every column of $p_j(v)$ lies in $U_j \cap u^\perp$. Since ι_j preserves rank,

$$\operatorname{rank}\bigl(\iota_j(p_j(v))\bigr) \leq k_j - 1, \quad \text{so } v \in \Sigma_j.$$

Conversely, let $v \in \Sigma_j$. Since $\iota_j(p_j(\Lambda))$ is commensurable with a rational lattice, the left kernel of $\iota_j(p_j(v))$ contains a nonzero rational covector. Pulling this covector back through a_j, and using the Euclidean identification of U_j^* with U_j, we obtain a nonzero $u \in U_j$ such that $\mathcal{L}(\mathbb{R}u)$ is $p_j(\Lambda)^*$-rational and orthogonal to $p_j(v)$. Because $p_j(\Lambda)^* = \Lambda^* \cap W_j$, the space $\mathcal{L}(\mathbb{R}u)$ is Λ^*-rational. Hence $V := \mathcal{L}(\mathbb{R}u)^\perp$ is a Λ-rational column-isotropic subspace of dimension $d(d-1)$, with $V^\perp \subset W_j$. Since $v - p_j(v) \in W_j^\perp$, the orthogonality of $p_j(v)$ to $\mathcal{L}(\mathbb{R}u)$ implies $v \in V$. This proves (10.14), and hence (10.12).

It remains to estimate the overlap of two distinct main pieces. By construction, $U_j \cap U_{j'} = \{0\}$ for $j \neq j'$. Put

$$W := W_j \oplus W_{j'} = \mathcal{L}(U_j \oplus U_{j'}).$$

The lattice $\Lambda^* \cap W$ is commensurable with $(\Lambda^* \cap W_j) \oplus (\Lambda^* \cap W_{j'})$. Using the two tensor-product coordinate maps on the two summands, the joint quotient lattice is commensurable with a full-rank rational lattice in $\mathrm{M}_{k_j,d}(\mathbb{R}) \oplus \mathrm{M}_{k_{j'},d}(\mathbb{R})$. Thus, after scaling, we may overcount its points by pairs of integral matrices. The number of pairs (A, A') of norm $O(T)$ satisfying

$$\operatorname{rank}(A) \leq k_j - 1, \qquad \operatorname{rank}(A') \leq k_{j'} - 1$$

is $O\bigl(T^{d(k_j-1)+d(k_{j'}-1)}\bigr)$, by Theorem 10.1; if one of the k's equals 1, the corresponding matrix is zero and contributes $O(1)$. The kernel of the joint quotient map has dimension $d(d-k_j-k_{j'})$, so each pair has $O(T^{d(d-k_j-k_{j'})})$ lifts in Λ. Therefore

$$\#\{v \in \Sigma_j \cap \Sigma_{j'} : \|v\| \leq T\} \ll T^{d(k_j-1)+d(k_{j'}-1)+d(d-k_j-k_{j'})} = T^{d(d-2)}.$$

This proves (10.13). The row statement follows after transposition. $\square$

Combining Lemma 10.6 with Lemma 10.7, we obtain constants $c^{\mathrm{sing}}_{\Lambda,\mathrm{col}}, c^{\mathrm{sing}}_{\Lambda,\mathrm{row}} \geq 0$ such that

$$\#\{v \in \Lambda^{(1)}_{\mathrm{iso}} : \|v\| \leq T\} \sim c^{\mathrm{sing}}_{\Lambda,\mathrm{col}} T^{d(d-1)}, \tag{10.15}$$

$$\#\{v \in \Lambda^{(2)}_{\mathrm{iso}} : \|v\| \leq T\} \sim c^{\mathrm{sing}}_{\Lambda,\mathrm{row}} T^{d(d-1)}. \tag{10.16}$$

Indeed, each main piece Σ_j has the asserted asymptotic, the finitely many lower-dimensional spaces contribute $O(T^{d(d-2)})$, and (10.13) gives the same bound for every pairwise overlap of main pieces.

10.5. **Mixed column–row intersections.** The column and row asymptotics must now be combined by inclusion–exclusion. The following compatibility lemma gives rational coordinates for the joint projection associated with one column piece and one row piece. We use the notation C and R for the two tensor factors.

Lemma 10.8 (Compatible split coordinates). *Let C and R be d-dimensional real vector spaces, let $U < \mathsf{C}$ and $V < \mathsf{R}$, and put*

$$W := U \otimes \mathsf{R}, \qquad Z := \mathsf{C} \otimes V.$$

Let $\Delta < \mathsf{C} \otimes \mathsf{R}$ be a lattice. Assume that W and Z are Δ^*-rational and that

$$\Delta^* \cap W \qquad \textit{and} \qquad \Delta^* \cap Z$$

are commensurable with tensor-product lattices in W and Z, respectively. Then one can choose bases of C and R, adapted to U and V, such that the joint quotient image of Δ in the three blocks

$$U \otimes V, \qquad U \otimes (\mathsf{R}/V), \qquad (\mathsf{C}/U) \otimes V$$

is a full-rank rational lattice.

Proof. Up to commensurability, write

$$\Delta^* \cap W \asymp L_U \otimes L_{\mathsf{R}}, \qquad \Delta^* \cap Z \asymp M_{\mathsf{C}} \otimes M_V,$$

where $L_U < U$, $L_{\mathsf{R}} < \mathsf{R}$, $M_{\mathsf{C}} < \mathsf{C}$, and $M_V < V$ are full lattices in the indicated spaces. Since $W \cap Z = U \otimes V$ is Δ^*-rational, $\Delta^* \cap (W \cap Z)$ is full in $W \cap Z$. It is commensurable both with

$$L_U \otimes (L_{\mathsf{R}} \cap V) \qquad \text{and with} \qquad (M_{\mathsf{C}} \cap U) \otimes M_V.$$

In particular, V is rational with respect to L_{R}, U is rational with respect to M_{C}, and the two displayed product lattices are commensurable.

We use the elementary fact that if two full tensor-product lattices $L_1 \otimes L_2$ and $L'_1 \otimes L'_2$ are commensurable, then there is a scalar $c \neq 0$ such that L_1 is commensurable with cL'_1 and L_2 is commensurable with $c^{-1}L'_2$. Equivalently, if a Kronecker product $A_1 \otimes A_2$ of two invertible change-of-basis matrices has rational entries, then, after multiplying A_1 by a scalar and A_2 by the inverse scalar, both A_1 and A_2 have rational entries.

It follows that the lattice structures induced on U by L_U and $M_{\mathsf{C}} \cap U$ are homothetically commensurable, and similarly the lattice structures induced on V by $L_{\mathsf{R}} \cap V$ and M_V are homothetically commensurable. Choose a basis of C adapted to the rational subspace U and the lattice M_{C}, and choose a basis of R adapted to V and the lattice L_{R}. Rescaling one tensor-factor basis absorbs the common homothety. In these coordinates, both $\Delta^* \cap W$ and $\Delta^* \cap Z$ are rational lattices.

The dual quotient lattices $p_W(\Delta)$ and $p_Z(\Delta)$ are therefore rational in compatible coordinates. Moreover, the kernel of the joint quotient map is $W^\perp \cap Z^\perp$, which is Δ-rational because $W + Z$ is Δ^*-rational. Hence the joint quotient image is a lattice; since its coordinate projections are rational, it is a full-rank rational lattice in the three displayed blocks. ☐

Lemma 10.9 (Mixed-piece estimate)**.** *Let Σ_{c} be one of the column pieces and Σ_{r} one of the row pieces supplied by Lemma 10.7. Then*

$$\#\{v \in \Sigma_{\mathrm{c}} \cap \Sigma_{\mathrm{r}} : \|v\| \le T\} \ll T^{d^2-d-1}(\log T)^{3d}.$$

Proof. Let k and ℓ be the dimensions of the tensor factors defining the column and row pieces, respectively. Thus $1 \le k, \ell \le d-1$. The dual subspaces defining these two pieces are Λ^*-rational, their intersections with Λ^* are tensor-product split, and their intersection is again Λ^*-rational. Hence Lemma 10.8 applies. After fixed changes of basis in the two tensor factors, the joint quotient image of Λ in the three relevant blocks is a rational lattice. Scaling only changes constants, so we may overcount this quotient by integral matrices.

The omitted fourth block is the kernel of the joint projection and has dimension $(d-k)(d-\ell)$. Consequently, each projected point has $O(T^{(d-k)(d-\ell)})$ lifts of norm at most T. We write the three quotient blocks in the form

$$\begin{pmatrix} A & B \\ C & * \end{pmatrix},$$

where the star denotes the omitted kernel block, and

$$A \in \mathrm{M}_{k,\ell}(\mathbb{Z}), \qquad B \in \mathrm{M}_{k,d-\ell}(\mathbb{Z}), \qquad C \in \mathrm{M}_{d-k,\ell}(\mathbb{Z}).$$

The two piece conditions become

$$\operatorname{rank}\begin{pmatrix} A & B \end{pmatrix} \le k-1, \qquad \operatorname{rank}\begin{pmatrix} A \\ C \end{pmatrix} \le \ell - 1. \tag{10.17}$$

Interchanging the column and row pieces if necessary, assume $k \ge \ell$.

We stratify by $r := \operatorname{rank}(A)$, where $0 \le r \le \ell - 1$. Applying Corollary 10.3 to $A^\mathsf{T} \in \mathrm{M}_{\ell,k}(\mathbb{Z})$ gives

$$\#\{A : \|A\| \le T,\ \operatorname{rank}(A) = r\} \ll T^{kr}(\log T)^d. \tag{10.18}$$

First suppose that $r \ge k + \ell - d$. Then $d - \ell \ge k - r$ and $d - k \ge \ell - r$. For fixed A, applying Lemma 10.4 to (A, B), and then its transposed version to (A, C), gives

$$\#\{B : \operatorname{rank}\begin{pmatrix} A & B \end{pmatrix} \le k-1\} \ll T^{(d-\ell)(k-1)}(\log T)^d,$$
$$\#\{C : \operatorname{rank}\begin{pmatrix} A \\ C \end{pmatrix} \le \ell-1\} \ll T^{(d-k)(\ell-1)}(\log T)^d.$$

Including the lifts in the omitted block, the rank-r stratum contributes at most

$$T^{kr+(d-\ell)(k-1)+(d-k)(\ell-1)+(d-k)(d-\ell)}(\log T)^{3d}.$$

Since $r \le \ell - 1$, the exponent of T is at most $d^2 - 2d + \ell \le d^2 - d - 1$.

Now suppose that $r < k + \ell - d$. Then the two rank conditions in (10.17) are automatic once A is fixed, because

$$r + (d - \ell) \le k - 1, \qquad r + (d - k) \le \ell - 1.$$

Thus B, C, and the omitted block contribute respectively

$$O(T^{k(d-\ell)}), \qquad O(T^{\ell(d-k)}), \qquad O(T^{(d-k)(d-\ell)}).$$

Together with (10.18), the exponent of T is

$$kr + k(d-\ell) + \ell(d-k) + (d-k)(d-\ell) = d^2 - k(\ell - r).$$

Since $r \le k + \ell - d - 1$, we have $\ell - r \ge d - k + 1$. Hence

$$d^2 - k(\ell - r) \le d^2 - k(d - k + 1) \le d^2 - d - 1.$$

The last inequality follows from

$$k(d - k + 1) - (d + 1) = (k - 1)(d - k) - 1 \ge 0;$$

this second case can occur only for $2 \le k \le d - 1$.

Summing over $0 \le r \le \ell - 1$ proves the lemma. □

10.6. **Proof of Theorem 3.15.**

Proof of Theorem 3.15. The column and row asymptotics are given by (10.15) and (10.16). We claim that their intersection is negligible.

Apply the finite piece decomposition in both the column and row cases. The intersection of a lower-dimensional piece with any other set contributes at most $O(T^{d(d-2)})$. There are only finitely many pairs of main column and row pieces, and Lemma 10.9 gives

$$\#\left\{v \in \Lambda_{\text{iso}}^{(1)} \cap \Lambda_{\text{iso}}^{(2)} : \|v\| \le T\right\} \ll T^{d^2-d-1}(\log T)^{3d} + T^{d(d-2)}.$$

In particular,

$$\#\left\{v \in \Lambda_{\text{iso}}^{(1)} \cap \Lambda_{\text{iso}}^{(2)} : \|v\| \le T\right\} = o(T^{d(d-1)}). \tag{10.19}$$

Using (10.11) and inclusion–exclusion, we therefore obtain

$$\begin{aligned}\#\{v \in \Lambda_{\text{iso}} : \|v\| \le T\} &= \#\{v \in \Lambda_{\text{iso}}^{(1)} : \|v\| \le T\} + \#\{v \in \Lambda_{\text{iso}}^{(2)} : \|v\| \le T\} \\ &\quad - \#\{v \in \Lambda_{\text{iso}}^{(1)} \cap \Lambda_{\text{iso}}^{(2)} : \|v\| \le T\} \\ &= \left(c_{\Lambda,\text{col}}^{\text{sing}} + c_{\Lambda,\text{row}}^{\text{sing}}\right) T^{d(d-1)} + o(T^{d(d-1)}).\end{aligned}$$

Thus the limit in Theorem 3.15 exists, with

$$c_\Lambda^{\rm sing} = c_{\Lambda,\rm col}^{\rm sing} + c_{\Lambda,\rm row}^{\rm sing} \geq 0.$$

This proof also shows that Theorem 3.15 applies to determinant-rational lattices that are not $\mathbb{Q}$-split. □

11. Fiber integrals and integral identities

In this section, we isolate the kernel governing the asymptotic K-averages that appear in the counting argument.

For $v \in \mathrm{M}_d(\mathbb{R})$, we write

$$v = \begin{pmatrix} v^0 & v^+ \\ v^- & v_{dd} \end{pmatrix}, \tag{11.1}$$

where $v^0 \in \mathrm{M}_{d-1}(\mathbb{R})$, $v^+ \in \mathbb{R}^{d-1}$, and $v^- \in (\mathbb{R}^{d-1})^\intercal$.

Let

$$\mathrm{M}_d^+ = \{v \in \mathrm{M}_d(\mathbb{R}) : \det v^0 > 0\}.$$

For a test function $f \in C_c(\mathrm{M}_d^+)$, the main object is the fiber integral $J_f(\mathsf{r}, \zeta)$ defined in (11.1), where $\mathsf{r} \in B_0^+$ records the $(d-1)$-dimensional singular-value data and ζ records the determinant. The section has two purposes. First, we show that compactly supported functions on $B_0^+ \times \mathbb{R}$ can be realized through the kernels J_f. Second, we establish the integral identities satisfied by J_f, which identify the limiting functional arising from the dynamical averaging. These identities will be the key input in the counting argument of the next section.

By the singular value decomposition, for each $v \in \mathrm{M}_d(\mathbb{R})$, there exists a unique diagonal element

$$b(v) := \mathrm{diag}\big(e^{\kappa_1(v)}, \ldots, e^{\kappa_{d-1}(v)}, \mathrm{sgn}(\det v) e^{\kappa_d(v)}\big)$$

with $\kappa_1(v) \geq \cdots \geq \kappa_d(v) \geq -\infty$ such that

$$v = k \cdot b(v)$$

for some $k \in K$. When $\det(v) = 0$, we have $e^{\kappa_i(v)} = 0$, i.e., $\kappa_i(v) = -\infty$ for some i.

For $v \in \mathrm{M}_d^+$, set

$$b^0(v) = \mathrm{diag}(e^{\kappa_1(v)}, \ldots, e^{\kappa_{d-1}(v)}) \in \mathrm{GL}_{d-1}^+(\mathbb{R})$$

where $\kappa_1(v) \geq \cdots \geq \kappa_{d-1}(v) > -\infty$. Note that the sign of $\det v$ is recorded in the last diagonal entry of $b(v)$; accordingly $b^0(v)$ lies in $\mathrm{GL}_{d-1}^+(\mathbb{R})$, while the full determinant sign is carried by the parameter $\zeta = \det v \in \mathbb{R}$.

11.1. **The fiber integrals J_f and approximation of K-averages.** Let

$$\begin{aligned} B^0_+ &= \{\mathsf{r} := \operatorname{diag}(r_1, \cdots, r_{d-1}) \in \mathrm{GL}^+_{d-1}(\mathbb{R}) : r_1 \geq \cdots \geq r_{d-1} > 0\} \\ &= \{b^0(v) : v \in \mathrm{M}^+_d\}. \end{aligned}$$

Note that $M = \mathrm{SO}(d-1) \times \mathrm{SO}(d-1)$ acts on $\mathrm{GL}^+_{d-1}(\mathbb{R})$ by

$$(m_1, m_2) \cdot A := m_1 A m_2.$$

By the Cartan decomposition of $\mathrm{GL}^+_{d-1}(\mathbb{R})$, for any $A \in \mathrm{GL}^+_{d-1}(\mathbb{R})$ there exists a unique $\mathsf{r} \in B^0_+$ such that $A \in M.\mathsf{r}$. Thus any continuous function on B^0_+ determines a unique continuous M-invariant function on $\mathrm{GL}^+_{d-1}(\mathbb{R})$, and vice versa.

Definition 11.1. Given $f \in C_c(\mathrm{M}^+_d)$, we define the associated fiber integral function

$$J_f : B^0_+ \times \mathbb{R} \to \mathbb{R}$$

as follows: for any $(\mathsf{r}, \zeta) \in B^0_+ \times \mathbb{R}$,

$$J_f(\mathsf{r}, \zeta) := \frac{1}{\det(\mathsf{r})^2} \int_{\mathbb{R}^{d-1} \times \mathbb{R}^{d-1}} \int_M f \begin{pmatrix} m.\mathsf{r} & x_2 \\ x_1^\intercal & x_{dd} \end{pmatrix} dm\, dx_1\, dx_2,$$

where $x_{dd} = x_{dd}(\mathsf{r}, \zeta, m, x_1, x_2)$ is determined by the condition

$$\det \begin{pmatrix} m.\mathsf{r} & x_2 \\ x_1^\intercal & x_{dd} \end{pmatrix} = \zeta.$$

We may consider J_f as a function on $\mathrm{GL}^+_{d-1}(\mathbb{R}) \times \mathbb{R}$:

$$J_f(A, \zeta) := J_f(\mathsf{r}, \zeta),$$

where $\mathsf{r} \in B^0_+$ is the unique element such that $A \in M.\mathsf{r}$.

Lemma 11.2. *For any M-invariant function $h \in C_c(\mathrm{GL}_{d-1}(\mathbb{R})^+ \times \mathbb{R})$, there exists $f \in C_c(\mathrm{M}^+_d)$ such that*

$$J_f(A, \zeta) = h(A, \zeta) \qquad \textit{for all } (A, \zeta) \in \mathrm{GL}_{d-1}(\mathbb{R})^+ \times \mathbb{R}.$$

Proof. Choose a nonnegative $\psi \in C_c(\mathbb{R}^{d-1} \times \mathbb{R}^{d-1})$ such that

$$\int_{\mathbb{R}^{d-1} \times \mathbb{R}^{d-1}} \psi(x_1, x_2)\, dx_1\, dx_2 = 1.$$

For an M-invariant function $h \in C_c(\mathrm{GL}_{d-1}(\mathbb{R})^+ \times \mathbb{R})$, define $f_h : \mathrm{M}^+_d \to \mathbb{R}$ by

$$f_h \begin{pmatrix} m \cdot \mathsf{r} & x_2 \\ x_1^\intercal & x_{dd} \end{pmatrix} := \det(\mathsf{r})^2\, h\Bigg(\mathsf{r}, \det \begin{pmatrix} m \cdot r & x_2 \\ x_1^\intercal & x_{dd} \end{pmatrix}\Bigg)\, \psi(x_1, x_2) \tag{11.2}$$

where $\mathsf{r} \in B^0_+$, $m \in M$, $x_1, x_2 \in \mathbb{R}^{d-1}$. Then $f \in C_c(\mathrm{M}^+_d)$, and by the definition of J_f and ψ,

$$J_f(\mathsf{r}, \zeta) = h(\mathsf{r}, \zeta) \quad \text{for all } \mathsf{r} \in B^0_+.$$

This proves the claim. ∎

We set

$$\mathbb{S}^{N-1} = \mathbb{S}(\mathrm{M}_d(\mathbb{R})) = \{v \in \mathrm{M}_d(\mathbb{R}) : \|v\| = 1\}.$$

We denote by $\widehat{v}$ the adjugate matrix $\mathrm{adj}(v)$, that is, the transpose of the cofactor matrix of v. For $v \in \mathrm{M}_d(\mathbb{R})$,

$$\widehat{v} \neq 0 \quad \Longleftrightarrow \quad \mathrm{rank}(v) \geq d-1.$$

Equivalently, $\widehat{v} = 0$ precisely on the rank-at-most-$(d-2)$ locus. Thus $\widehat{v}/\|\widehat{v}\|$ is defined exactly on the regular part of the determinant hypersurface together with the full-rank locus. We compare the K-average of $f(a_t k \cdot v)$ with the corresponding averaged kernel J_f.

Proposition 11.3. *Let $f \in C_c(\mathrm{M}_d^+)$ and $\nu \in C(\mathbb{S}^{N-1})$ be nonnegative functions. For any $\varepsilon > 0$, there exist constants $t_0, T_0 > 0$ such that the following holds uniformly for all $t \geq t_0$ and all $v \in \mathrm{M}_d(\mathbb{R})$ with $\|v\| \geq T_0$ and $\widehat{v} \neq 0$,*

$$\left| \omega_{d-1}^2 \, e^{2d(d-1)t} \int_K f(a_t k \cdot v) \, \nu\big((k^\intercal \cdot E_{dd})^\intercal\big) \, dk \right. \\ \left. - J_f\big(b^0(e^{-2t} v), \det(v)\big) \, \nu\Big(\frac{\widehat{v}}{\|\widehat{v}\|}\Big) \right| < \varepsilon$$

where ω_{d-1} denotes the surface volume of unit sphere $\{x_1^2 + \cdots + x_d^2 = 1\}$.

Proof. For simplicity, we write $b_v = b(v)$. We write

$$I_t(v) := \int_K f(a_t k \cdot v) \, \nu\big((k^\intercal \cdot E_{dd})^\intercal\big) \, dk.$$

For $k = (k_1, k_2) \in K$, define

$$k_* := (k_2^\intercal, k_1^\intercal) \in K.$$

Then

$$(k^\intercal \cdot E_{dd})^\intercal = k_* \cdot E_{dd},$$

and, since $\widehat{k_i} = k_i^{-1} = k_i^\intercal$ for $k_i \in \mathrm{SO}(d)$, we have

$$k_* \cdot \widehat{k \cdot u} = \widehat{u} \qquad \text{for all } u \in \mathrm{M}_d(\mathbb{R}).$$

Using Haar invariance and the change of variables $k \mapsto k k_v^{-1}$, we obtain

$$I_t(v) = \int_K f(a_t k \cdot b_v) \, \nu\big((k k_v^{-1})_* \cdot E_{dd}\big) \, dk.$$

If $a_t k \cdot b_v \notin \mathrm{supp}(f)$ for all $k \in K$, then the left-hand side in the statement vanishes. Since the term $J_f\big(b^0(e^{-2t} v), \det(v)\big)$ also vanishes for t sufficiently large, there is nothing to prove in this case. We may therefore assume that there exists $k \in K$ such that

$$a_t k \cdot b_v \in \mathrm{supp}(f).$$

Step 1: localization near E_{dd}. Since f is compactly supported in M_d^+, there exists $c > 0$ such that

$$\det(u^0) \geq c \qquad \text{for all } u \in \operatorname{supp}(f).$$

Now

$$a_t k \cdot b_v = \begin{pmatrix} e^{-2t}(k \cdot b_v)^0 & e^{-(d-2)t}(k \cdot b_v)^+ \\ e^{-(d-2)t}(k \cdot b_v)^- & e^{2(d-1)t}(k \cdot b_v)_{dd} \end{pmatrix},$$

so the support condition implies

$$\det(a_t k \cdot b_v)^0 = e^{-2(d-1)t} \det((k \cdot b_v)^0) \geq c.$$

Hence

$$|(\widehat{k \cdot b_v})_{dd}| = |\det((k \cdot b_v)^0)| \geq c\, e^{2(d-1)t}.$$

On the other hand, any other entry of $\widehat{k \cdot b_v}$ is $O(e^{(d-2)t})$, because it is a $(d-1) \times (d-1)$-minor involving at least one row or column from the lower-right corner of $a_t k \cdot b_v$. Therefore

$$\left\| \frac{\widehat{k \cdot b_v}}{\|\widehat{k \cdot b_v}\|} - E_{dd} \right\| = O(e^{-dt}). \tag{11.3}$$

By uniform continuity of ν, given $\delta > 0$, there exists t_0 such that whenever $t \geq t_0$ and $a_t k \cdot b_v \in \operatorname{supp}(f)$, we have

$$\left| \nu\big((kk_v^{-1})_* \cdot E_{dd}\big) - \nu\left((kk_v^{-1})_* \cdot \frac{\widehat{k \cdot b_v}}{\|\widehat{k \cdot b_v}\|}\right) \right| < \delta.$$

Using the covariance of the adjugate, this becomes

$$\left| \nu\big((kk_v^{-1})_* \cdot E_{dd}\big) - \nu\left(\frac{\widehat{v}}{\|\widehat{v}\|}\right) \right| < \delta. \tag{11.4}$$

Step 2: local parametrization of K. For $1 \leq i \leq d-1$ and $\theta \in \mathbb{R}$, define

$$\mathsf{n}_i(\theta) = \begin{pmatrix} I_{i-1} & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & -\sin\theta \\ 0 & 0 & I_{d-i-1} & 0 \\ 0 & \sin\theta & 0 & \cos\theta \end{pmatrix} \in \mathrm{SO}(d),$$

and set

$$\mathsf{n}(\theta_1, \ldots, \theta_{d-1}) := \mathsf{n}_1(\theta_1) \cdots \mathsf{n}_{d-1}(\theta_{d-1}).$$

Let

$$\mathsf{N} := \{\mathsf{n}(\theta_1, \ldots, \theta_{d-1}) : \ \theta_i \in (-\tfrac{\pi}{2}, \tfrac{\pi}{2})\}.$$

If $a_t k \cdot b_v \in \operatorname{supp}(f)$, then by (11.3) we have

$$\|k_* \cdot E_{dd} - E_{dd}\| = O(e^{-dt}).$$

Writing $k = (k_1, k_2) \in \mathrm{SO}(d) \times \mathrm{SO}(d)$, this means that both k_1 and k_2 are $O(e^{-dt})$-close to the subgroup $\mathrm{SO}(d-1)$ stabilizing e_d. Since the projection

$$\mathsf{N} \to \mathrm{SO}(d-1) \backslash \mathrm{SO}(d)$$

is a local diffeomorphism at the identity, we may write

$$k_1 = m_1\,\mathsf{n}(\theta_1^-,\dots,\theta_{d-1}^-), \qquad k_2 = m_2\,\mathsf{n}(\theta_1^+,\dots,\theta_{d-1}^+)$$

with $m_1, m_2 \in \mathrm{SO}(d-1)$ and

$$\theta_i^{\pm} = O(e^{-dt}) \qquad (1 \le i \le d-1).$$

Thus

$$k = m(\mathsf{n}^-, (\mathsf{n}^+)^\intercal), \qquad m := (m_1, m_2) \in M,$$

where

$$\mathsf{n}^{\pm} := \mathsf{n}(\theta_1^{\pm},\dots,\theta_{d-1}^{\pm}).$$

Step 3: approximation by the J_f-kernel. Since $m \in M$ commutes with a_t, we have

$$a_t k \cdot b_v = m \cdot \Big(a_t(\mathsf{n}^-, (\mathsf{n}^+)^\intercal) a_{-t} \cdot (a_t \cdot b_v)\Big).$$

For $x \in \mathbb{R}$, define

$$n_i^-(x) := \begin{pmatrix} I_{i-1} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & I_{d-i-1} & 0 \\ 0 & x & 0 & 1 \end{pmatrix}, \qquad n_i^+(x) := (n_i^-(x))^\intercal.$$

Set

$$x_\theta := e^{dt}\sin\theta.$$

Since $\theta = O(e^{-dt})$, we have

$$\cos\theta = 1 + O(e^{-2dt}), \qquad e^{-dt}\sin\theta = O(e^{-2dt}),$$

and therefore

$$b_t \mathsf{n}_i(\theta) b_{-t} = n_i^-(x_\theta) + O(e^{-dt}), \qquad b_t \mathsf{n}_i(\theta)^\intercal b_{-t} = n_i^+(x_\theta) + O(e^{-dt}).$$

Now for each $1 \le i \le d-1$, define

$$x_i^- := e^{(d-2)t} e^{\kappa_i(v)} \sin\theta_i^-, \qquad x_i^+ := e^{(d-2)t} e^{\kappa_i(v)} \sin\theta_i^+,$$

and let

$$x^- = (x_1^-,\dots,x_{d-1}^-)^\intercal, \qquad x^+ = (x_1^+,\dots,x_{d-1}^+)^\intercal.$$

Then $a_t k \cdot b_v$ is approximated by

$$\mathbf{x}(m, x^-, x^+) := \begin{pmatrix} m \cdot b^0(e^{-2t} v) & x^+ \\ (x^-)^\intercal & x_{dd} \end{pmatrix},$$

up to an error $O(e^{-dt})$, where x_{dd} is uniquely chosen so that

$$\det \mathbf{x}(m, x^-, x^+) = \det(v).$$

Since $\det(\mathbf{x}^0)$ is bounded below and the support of f is compact, we also have

$$|x_{dd} - (a_t k \cdot b_v)_{dd}| = O(e^{-dt}).$$

By uniform continuity of f, after enlarging t_0 if necessary we may assume

$$|f(a_t k \cdot b_v) - f(\mathbf{x}(m, x^-, x^+))| < \delta \qquad (t \ge t_0). \tag{11.5}$$

Step 4: the Jacobian and the appearance of ω_{d-1}. For each $1 \le i \le d-1$,

$$dx_i^{\pm} = e^{(d-2)t} e^{\kappa_i(v)} \cos\theta_i^{\pm}\, d\theta_i^{\pm}.$$

Hence

$$dx^-\, dx^+ = e^{2(d-2)(d-1)t} \det(b^0(v))^2 \prod_{i=1}^{d-1} \cos\theta_i^- \; \cos\theta_i^+\, d\theta^-\, d\theta^+.$$

We now recall the standard factorization of Haar probability measure on SO(d). If dk_i denotes Haar probability measure on SO(d) and dm_i Haar probability measure on SO($d-1$), then in the coordinates $k_i = m_i \mathsf{n}(\theta^{\pm})$ we have

$$dk_i = \omega_{d-1}^{-1} \left(\prod_{j=1}^{d-1} (\cos\theta_j^{\pm})^{j-1} \right) dm_i\, d\theta_1^{\pm} \cdots d\theta_{d-1}^{\pm}.$$

Here

$$\omega_{d-1} = \mathrm{vol}(\mathbb{S}^{d-1} \subset \mathbb{R}^d)$$

is exactly the normalizing constant coming from the standard spherical coordinate formula on the unit sphere. Since $dk = dk_1\, dk_2$ and $dm = dm_1\, dm_2$, this gives

$$dk = \omega_{d-1}^{-2} \left(\prod_{j=1}^{d-1} (\cos\theta_j^-)^{j-1} \right) \left(\prod_{j=1}^{d-1} (\cos\theta_j^+)^{j-1} \right) dm\, d\theta^-\, d\theta^+.$$

Equivalently,

$$dm\, d\theta^-\, d\theta^+ = \omega_{d-1}^2 \left(\prod_{j=1}^{d-1} (\cos\theta_j^-)^{-(j-1)} \right) \left(\prod_{j=1}^{d-1} (\cos\theta_j^+)^{-(j-1)} \right) dk.$$

Combining this with the formula for $dx^- dx^+$, we obtain

$$\begin{aligned} dm\, dx^-\, dx^+ = {} & \omega_{d-1}^2 e^{2(d-2)(d-1)t} \det(b^0(v))^2 \\ & \cdot \left(\prod_{j=1}^{d-1} (\cos\theta_j^-)^{2-j} \right) \left(\prod_{j=1}^{d-1} (\cos\theta_j^+)^{2-j} \right) dk. \end{aligned}$$

Since $\theta_j^{\pm} = O(e^{-dt})$, any cosine factor equals $1 + O(e^{-dt})$, and therefore

$$dm\, dx^-\, dx^+ = \big(1 + O(e^{-dt})\big)\, \omega_{d-1}^2 e^{2(d-2)(d-1)t} \det(b^0(v))^2\, dk.$$

Finally, since

$$b^0(e^{-2t} v) = e^{-2t} b^0(v),$$

we have

$$e^{2(d-2)(d-1)t} \det(b^0(v))^2 = e^{2d(d-1)t} \det\big(b^0(e^{-2t} v)\big)^2.$$

Thus

$$dm\, dx^-\, dx^+ = \big(1 + O(e^{-dt})\big)\, \omega_{d-1}^2 e^{2d(d-1)t} \det\big(b^0(e^{-2t} v)\big)^2\, dk. \tag{11.6}$$

Step 5: conclusion. Using (11.4), (11.5), and (11.6), we obtain

$$\omega_{d-1}^2 e^{2d(d-1)t} \int_K f(a_t k \cdot v)\, \nu\big((k^\intercal \cdot E_{dd})^\intercal\big)\, dk$$
$$= \nu\Big(\frac{\widehat{v}}{\|\widehat{v}\|}\Big) \frac{1}{\det(b^0(e^{-2t}v))^2} \int_{\mathbb{R}^{d-1}\times\mathbb{R}^{d-1}} \int_M f(\mathbf{x}(m, x^-, x^+))\, dm\, dx^-\, dx^+ + O(\delta).$$

By definition of J_f, the inner integral is exactly $J_f\big(b^0(e^{-2t}v), \det(v)\big)$. Hence

$$\left|\omega_{d-1}^2 e^{2d(d-1)t} \int_K f(a_t k \cdot v)\, \nu\big((k^\intercal \cdot E_{dd})^\intercal\big)\, dk - J_f\big(b^0(e^{-2t}v), \det(v)\big)\nu\Big(\frac{\widehat{v}}{\|\widehat{v}\|}\Big)\right| \ll \delta.$$

Choosing δ sufficiently small completes the proof. □

11.2. **Integral identities for J_f.** We now turn to the integral identities satisfied by J_f. These will convert the asymptotic kernel appearing in Proposition 11.3 into the explicit main term used in the counting argument.

Proposition 11.4. *Let $f \in C_c(\mathrm{M}_d^+)$ and $\nu \in C(\mathbb{S}^{N-1})$ be nonnegative. Then*

$$\lim_{t\to\infty} e^{-2d(d-1)t} \int_{\mathrm{M}_d(\mathbb{R})} J_f(b^0(e^{-2t}v), \det(v))\ \nu\Big(\frac{\widehat{v}}{\|\widehat{v}\|}\Big)\, dv$$
$$= \omega_{d-1}^2 \Big(\int_{\mathrm{M}_d(\mathbb{R})} f(v)\, dv\Big) \Big(\int_K \nu\big((k^\intercal \cdot E_{dd})^\intercal\big)\, dk\Big).$$

We will deduce this proposition from Lemmas 11.5 and 11.6 below. The following shows that J_f recovers the ambient Lebesgue integral of f.

Let $d^\times A := \det(A)^{-(d-1)}\, dA$ denote the Haar measure on $\mathrm{GL}_{d-1}^+(\mathbb{R})$, and let $d\mathsf{r}$ be the measure on B_+^0 characterized by

$$d^\times A = dm\, d\mathsf{r} \qquad \text{for } A = m.\mathsf{r}. \tag{11.7}$$

So if $\mathsf{r} = \mathrm{diag}(e^{a_1}, \cdots, e^{a_{d-1}})$, then $d\mathsf{r} = c\prod_{1\le i<j\le d-1} \sinh(a_i - a_j) da_1 \cdots da_{d-1}$.

Lemma 11.5. *For any $f \in C_c(\mathrm{M}_d^+)$, we have*

$$\int_{\mathrm{M}_d(\mathbb{R})} f\, dv = \int_{\mathbb{R}} \int_{\mathsf{r}\in B_+^0} J_f(\mathsf{r}, \zeta) \det(\mathsf{r})^d d\mathsf{r}\, d\zeta.$$

Proof. Since f is supported in M_d^+, we may write

$$\int_{\mathrm{M}_d(\mathbb{R})} f(v)\, dv = \int_{\mathrm{M}_d^+} f(v)\, dv.$$

Write

$$v = \begin{pmatrix} A & x_2 \\ x_1^\intercal & x_{dd} \end{pmatrix}, \qquad A \in \mathrm{GL}_{d-1}^+(\mathbb{R}), \quad x_1, x_2 \in \mathbb{R}^{d-1}, \quad x_{dd} \in \mathbb{R}.$$

Set

$$\zeta := \det(v) = \det(A)\big(x_{dd} - x_1^\intercal A^{-1} x_2\big).$$

For fixed A, x_1, x_2, ζ, this determines x_{dd} uniquely as

$$x_{dd}(A, \zeta, x_1, x_2) = \det(A)^{-1}\zeta + x_1^\intercal A^{-1} x_2.$$

Since $\frac{\partial \zeta}{\partial x_{dd}} = \det(A)$, the change of variables $(A, x_1, x_2, x_{dd}) \mapsto (A, x_1, x_2, \zeta)$ has Jacobian $\det(A)$, and therefore

$$dv = dA\, dx_1\, dx_2\, dx_{dd} = \frac{1}{\det(A)}\, dA\, dx_1\, dx_2\, d\zeta.$$

Hence

$$\int_{\mathrm{M}_d^+} f(v)\, dv = \int_{\mathbb{R}} \int_{\mathrm{GL}_{d-1}^+(\mathbb{R})} \int_{\mathbb{R}^{d-1}\times\mathbb{R}^{d-1}} f\begin{pmatrix} A & x_2 \\ x_1^\intercal & x_{dd}(A, \zeta, x_1, x_2) \end{pmatrix} \frac{dx_1\, dx_2\, dA}{\det(A)}\, d\zeta.$$

Now write

$$\frac{dA}{\det(A)} = \det(A)^{d-2}\, d^\times A.$$

Using (11.7), we get

$$\int_{\mathrm{M}_d^+} f(v)\, dv = \int_{\mathbb{R}} \int_{\mathsf{r}\in B_0^+} \int_{m\in M} \int_{\mathbb{R}^{d-1}\times\mathbb{R}^{d-1}} f\begin{pmatrix} m\cdot \mathsf{r} & x_2 \\ x_1^\intercal & x_{dd}(m\cdot \mathsf{r}, \zeta, x_1, x_2) \end{pmatrix} \cdot \det(\mathsf{r})^{d-2}\, dx_1\, dx_2\, dm\, d\mathsf{r}\, d\zeta.$$

Since

$$\int_{m\in M} \int_{\mathbb{R}^{d-1}\times\mathbb{R}^{d-1}} f\begin{pmatrix} m\cdot \mathsf{r} & x_2 \\ x_1^\intercal & x_{dd}(m\cdot r, \zeta, x_1, x_2) \end{pmatrix} dx_1\, dx_2\, dm = \det(\mathsf{r})^2\, J_f(\mathsf{r}, \zeta),$$

substituting this into the previous formula, we obtain

$$\int_{\mathrm{M}_d(\mathbb{R})} f(v)\, dv = \int_{\mathbb{R}} \int_{r\in B_0^+} J_f(r, \zeta)\, \det(\mathsf{r})^d\, d\mathsf{r}\, d\zeta,$$

as claimed. □

The next lemma gives the limiting integral identity in an abstract form, for a general test function h on $\mathrm{GL}_{d-1}^+(\mathbb{R}) \times \mathbb{R}$. Its role is to convert the asymptotic K-average obtained in the approximation step into an explicit volume integral; later we will apply it with $h = J_f$.

Lemma 11.6. *Let $h \in C_c(\mathrm{GL}_{d-1}^+(\mathbb{R}) \times \mathbb{R})$ be an M-invariant function and $\nu \in C(\mathbb{S}^{N-1})$ be nonnegative. Then*

$$\lim_{t\to\infty} e^{-2d(d-1)t} \int_{\mathrm{M}_d(\mathbb{R})} h(b^0(e^{-2t}v), \det(v))\nu\Big(\frac{\widehat{v}}{\|\widehat{v}\|}\Big)\, dv$$
$$= \omega_{d-1}^2 \left(\int_K \nu\big((k^\intercal \cdot E_{dd})^\intercal\big)\, dk\right)\left(\int_{\mathbb{R}} \int_{B_+^0} h(\mathsf{r}, \zeta)\, \det(\mathsf{r})^d d\mathsf{r}\, d\zeta\right).$$

Proof. Considering h as a function on $B_+^0 \times \mathbb{R}$, let $f_h \in C_c(\mathrm{M}_d^+)$ be as in (11.2). Apply Lemma 11.3 to f_h and ν. Let

$$\Xi_t := \left\{v \in \mathrm{M}_d(\mathbb{R}) : h(b^0(e^{-2t}v), \det v) \neq 0 \text{ or } \exists\, k \in K : \ f_h(a_t k\cdot v) \neq 0\right\}.$$

Since h and f_h are compactly supported, there exists $C > 1$ such that for all sufficiently large t,

$$\Xi_t \subset \left\{v \in \mathrm{M}_d(\mathbb{R}) : C^{-1}e^{2t} \leq e^{\kappa_i(v)} \leq Ce^{2t}\ (1 \leq i \leq d-1),\ |\det v| \leq C\right\}.$$

By the Cartan decomposition of $\mathrm{GL}_d^+(\mathbb{R})$ and its negative-determinant analogue, this implies

$$e^{-2d(d-1)t}\ \mathrm{vol}(\Xi_t) \ll 1. \tag{11.8}$$

Now choose $\varepsilon > 0$, and let $\varepsilon_1 > 0$ be so small that the right-hand side of (11.8) times ε_1 is $< \varepsilon$. By Lemma 11.3, there exist $t_0, T_0 > 0$ such that for all $t \geq t_0$ and all $v \in \Xi_t$ we have

$$\left| h(b^0(e^{-2t}v), \det(v))\, \nu\Big(\frac{\widehat{v}}{\|\widehat{v}\|}\Big) - \omega_{d-1}^2 e^{2d(d-1)t} \int_K f_h(a_t k \cdot v)\, \nu\big((k^\intercal \cdot E_{dd})^\intercal\big)\, dk \right| < \varepsilon_1.$$

Integrating this over Ξ_t, multiplying by $e^{-2d(d-1)t}$, and using (11.8), we obtain

$$\Bigg| e^{-2d(d-1)t} \int_{\mathrm{M}_d(\mathbb{R})} h(b^0(e^{-2t}v), \det(v))\, \nu\Big(\frac{\widehat{v}}{\|\widehat{v}\|}\Big)\, dv \\ - \omega_{d-1}^2 \int_{\mathrm{M}_d(\mathbb{R})} \int_K f_h(a_t k \cdot v)\, \nu\big((k^\intercal \cdot E_{dd})^\intercal\big)\, dk\, dv \Bigg| < \varepsilon.$$

Since the integrand vanishes outside Ξ_t, we may drop the restriction to Ξ_t. Now apply Fubini and the change of variables $u = a_t k \cdot v$. Since

$$\int_{\mathrm{M}_d(\mathbb{R})} f_h(a_t k \cdot v)\, dv = \int_{\mathrm{M}_d(\mathbb{R})} f_h(u)\, du,$$

we have

$$\omega_{d-1}^2 \int_{\mathrm{M}_d(\mathbb{R})} \int_K f_h(a_t k \cdot v)\, \nu\big((k^\intercal \cdot E_{dd})^\intercal\big)\, dk\, dv \\ = \omega_{d-1}^2 \Big(\int_K \nu\big((k^\intercal \cdot E_{dd})^\intercal\big)\, dk \Big) \Big(\int_{\mathrm{M}_d(\mathbb{R})} f_h(u)\, du \Big).$$

Finally, by Lemma 11.5 and the identity $J_{f_h} = h$,

$$\int_{\mathrm{M}_d(\mathbb{R})} f_h(u)\, du = \int_{\mathbb{R}} \int_{B_+^0} h(\mathsf{r}, \zeta)\ \det(\mathsf{r})^d\, d\mathsf{r}\, d\zeta.$$

Since $\varepsilon > 0$ was arbitrary, the result follows. □

Proof of Proposition 11.4. Apply Lemma 11.6 with $h = J_f$. The left-hand side is exactly the one appearing in Lemma 11.6. On the right-hand side we obtain

$$\omega_{d-1}^2 \Big(\int_K \nu\big((k^\intercal \cdot E_{dd})^\intercal\big)\, dk \Big) \Big(\int_{\mathbb{R}} \int_{B_+^0} J_f(\mathsf{r}, \zeta)\ \det(\mathsf{r})^d\, d\mathsf{r}\, d\zeta \Big).$$

By Lemma 11.5, the second factor equals $\int_{\mathrm{M}_d(\mathbb{R})} f(v)\, dv$. This proves the claim. □

We also record the following corollary of Proposition 11.4 for a later volume computation.

Corollary 11.7. *For any function $F \in C_c(\mathrm{M}_d^+ \times \mathbb{R})$ such that $F(k.v,\zeta) = F(v,\zeta)$ for any $k \in K$, we have*

$$\lim_{T\to\infty} T^{-d(d-1)} \int_{\mathrm{M}_d(\mathbb{R})} F(T^{-1}v, \det v)\, dv = \omega_{d-1}^2 \int_{\mathrm{M}_d(\mathbb{R})\times\mathbb{R}} F(v,\zeta)\, dv\, d\zeta.$$

Proof. By Lemma 11.2, it suffices to consider functions of the form

$$F(v,\zeta) = J_h(v^0,\zeta)\, \nu\left(\frac{\widehat{v}}{\|\widehat{v}\|}\right),$$

for some $h \in C_c(B_+^0 \times \mathbb{R})$. Then the claim is exactly Proposition 11.4. □

Remark 11.8 (Uniformity on compact singular-value windows). Let $\Omega \subset B_0^+ \times \mathbb{R}$ be compact. Then the parameters r, r^{-1}, and ζ are uniformly bounded on Ω. Consequently, all norm-comparison factors and weight factors appearing in the fiber-coordinate formulas of this section are uniformly bounded on Ω. Moreover, any such factor which is continuous and nonzero on the region under consideration is bounded away from zero on Ω.

We shall use this compactness uniformity without further comment when the (r,ζ)-parameter is restricted to a compact singular-value window.

12. Modified Siegel transforms and counting

In this section we convert the uniform height estimates proved above into the counting asymptotics. There are three distinct steps. First, we extend Shah's equidistribution theorem from bounded functions on the space of lattices to the modified Siegel transforms used here. Second, we compare the resulting dynamical averages with the fiber kernels from Section 11. Third, we pass from compact subsets of the singular-value parameter space to norm balls. The last step requires a separate estimate for the deep singular-value cusp.

Throughout the section, let $\Lambda \in X$ be a Diophantine lattice. Fix parameters $0 < \eta < \eta_0 < 1$ and $M > 1$ such that Λ is (η_0, M)-Diophantine. By Lemma 9.1, the (η, M)-quasi-null rational subspaces are precisely the rational isotropic subspaces. We use the all-degree modified height $\widehat{\alpha}_{\eta,M}$ from Section 6; thus, in every exterior degree, rational subspaces contained in a rational isotropic critical subspace are omitted.

12.1. **Modified Siegel transforms and moment truncation.**

For a lattice Δ, put

$$\Delta_{\mathrm{ni}} := \Delta - \Delta_{\mathrm{iso}}, \tag{12.1}$$

where Δ_{iso} is the union of all proper Δ-rational column- and row-isotropic subspaces. If $f \in C_c(\mathrm{M}_d(\mathbb{R}))$, define

$$\widetilde{f}_{\mathrm{ni}}(\Delta) := \sum_{v\in\Delta_{\mathrm{ni}}} f(v).$$

Equivalently,

$$\widetilde{f}_{\mathrm{ni}}(h;\Delta) := \widetilde{f}_{\mathrm{ni}}(h\Delta) = \sum_{v\in\Delta_{\mathrm{ni}}} f(hv), \qquad h \in H,$$

because the H-action preserves the column- and row-isotropic families.

We first record the counting form of the modified Lipschitz principle. The second part, in which one side of the box is allowed to grow, will be used in the singular-value shell estimates.

Lemma 12.1 (Modified Lipschitz principle)**.** *Let $\Delta \in X$, and let $\eta, M > 0$. Suppose that all (η, M)-quasi-null subspaces of Δ are isotropic. Then the following statements hold.*

(i) *For every nonnegative $f \in C_c(\mathrm{M}_d(\mathbb{R}))$,*

$$\widetilde{f}_{\mathrm{ni}}(h;\Delta) \ll_f 1 + \widehat{\alpha}_{\eta,M}(h;\Delta) \qquad (h \in H).$$

(ii) *Fix $R \geq 1$. For $L \geq 1$, let*

$$\mathscr{B}_R(L) := \{x = (x_{ij}) \in \mathrm{M}_d(\mathbb{R}) : |x_{ij}| \leq R \textit{ for all } (i,j) \neq (d,d), |x_{dd}| \leq RL\}.$$

Then

$$\#\{v \in \Delta_{\mathrm{ni}} : hv \in \mathscr{B}_R(L)\} \ll L\big(1 + \widehat{\alpha}_{\eta,M}(h;\Delta)\big)$$

where the implied constant is uniform for all $h \in H$ and $L \geq 1$.

Proof. We use the standard successive-minima proof of the Lipschitz principle; see [11, 13]. We indicate why the modified height is the correct height in the present setting.

Let $\mathscr{B}$ be either a fixed bounded set containing $\mathrm{supp}(f)$, or the box $\mathscr{B}_R(L)$, and consider

$$\mathscr{S} := h\Delta_{\mathrm{ni}} \cap \mathscr{B}.$$

In the successive-minima argument one forms a flag of rational subspaces by successively adjoining elements of $\mathscr{S}$. Every nonzero subspace in this flag is spanned by vectors of $h\Delta_{\mathrm{ni}}$, and therefore cannot be contained in a rational isotropic critical subspace. Consequently every subspace in the flag is among the admissible subspaces in the definition of $\widehat{\alpha}_{\eta,M}(e;h\Delta)$.

The usual flag estimate bounds $\#\mathscr{S}$ by a dimensional constant times

$$1 + \max_V \frac{\mathrm{vol}_{\dim V}(p_V\mathscr{B})}{\mathrm{covol}_V(h\Delta \cap V)},$$

where the maximum is taken over the subspaces occurring in the flag and p_V denotes orthogonal projection. For a fixed bounded set, all the projection volumes in the numerator are uniformly bounded, proving (i). For $\mathscr{B}_R(L)$, every positive-dimensional orthogonal projection has volume $O_R(L)$, since only one coordinate length depends on L. This proves (ii). ☐

We shall repeatedly use the following elementary uniform-integrability principle. It is stated for finite measures because the weighted K-averages appearing below need not be probability measures.

Lemma 12.2 (Moment truncation)**.** *Let Y be a locally compact second-countable Hausdorff space, let μ_j and μ be finite Borel measures on Y, and suppose that $\mu_j \to \mu$ weakly. Let $\Phi : Y \to \mathbb{R}$ be Borel measurable and continuous outside a μ-null set. If, for some $\beta > 1$,*

$$\sup_j \int_Y |\Phi|^\beta \, d\mu_j < \infty, \tag{12.2}$$

then $\Phi \in L^1(\mu)$ and

$$\lim_{j\to\infty} \int_Y \Phi \, d\mu_j = \int_Y \Phi \, d\mu.$$

Proof. For $A > 0$, let

$$\Phi_A := \max\{-A, \min\{\Phi, A\}\}.$$

The discontinuity set of Φ_A is contained in the discontinuity set of Φ, hence is μ-null. By the Portmanteau theorem in the form for bounded μ-a.e. continuous functions,

$$\lim_{j\to\infty} \int_Y \Phi_A \, d\mu_j = \int_Y \Phi_A \, d\mu.$$

The same applies to $\min\{|\Phi|^\beta, A\}$. Letting $A \to \infty$ and using monotone convergence gives

$$\int_Y |\Phi|^\beta \, d\mu \le \sup_j \int_Y |\Phi|^\beta \, d\mu_j.$$

In particular, $\Phi \in L^1(\mu)$. Moreover,

$$\int_{\{|\Phi|>A\}} |\Phi| \, d\mu_j \le A^{1-\beta} \int_Y |\Phi|^\beta \, d\mu_j,$$

and the same estimate holds for μ. Hence the tails are uniformly negligible. Therefore,

$$\left| \int_Y \Phi \, d\mu_j - \int_Y \Phi \, d\mu \right| \le \left| \int_Y \Phi_A \, d\mu_j - \int_Y \Phi_A \, d\mu \right| + \int_{\{|\Phi|>A\}} |\Phi| \, d\mu_j + \int_{\{|\Phi|>A\}} |\Phi| \, d\mu.$$

Letting first $j \to \infty$ and then $A \to \infty$ proves the claim. □

The next observation verifies the almost-everywhere continuity required in Lemma 12.2.

Lemma 12.3 (Regularity of the modified Siegel transform)**.** *For every $f \in C_c(\mathrm{M}_d(\mathbb{R}))$, the function $\widetilde{f}_{\mathrm{ni}} : X \to \mathbb{R}$ is Borel measurable. Moreover, with respect to the Haar measure m_X, it agrees almost everywhere with the ordinary Siegel transform $\widetilde{f}$. In particular, if $0 \notin \mathrm{supp}(f)$, then $\widetilde{f}_{\mathrm{ni}}$ is continuous outside an m_X-null set.*

Proof. Every rational subspace of a lattice $g\mathbb{Z}^N$ has the form gW, where $W < \mathbb{Q}^N$ is rational. For each fixed rational W, the condition that gW be column- or row-isotropic is a proper algebraic condition on g. There are only countably many rational subspaces W. Hence the set of lattices possessing a nonzero rational isotropic subspace is contained in a countable union of proper algebraic subvarieties of G, and therefore has Haar measure zero in X.

Outside this null set one has $\Delta_{\mathrm{ni}} = \Delta - \{0\}$, so the modified transform agrees with the ordinary Siegel transform. The latter is continuous whenever the support of f avoids the origin. Borel measurability follows directly by writing the sum in local lattice charts, or equivalently by expressing the exceptional condition as the countable union of the algebraic conditions above. $\square$

12.2. **Counting singular-value shells.** We now control the parts of the norm ball which are not covered by compact singular-value windows. The dynamical equidistribution input used above is only along the principal ray a_t; the remaining singular-value ratios are handled here by elementary shell estimates.

For a matrix v, let

$$\kappa_1(v) \geq \cdots \geq \kappa_d(v)$$

denote the logarithms of its singular values, with the convention $\kappa_i(v) = -\infty$ when the corresponding singular value is zero. Let

$$\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_{d-1}), \qquad \kappa_1 \geq \cdots \geq \kappa_{d-1}, \qquad \kappa_1 \geq 0.$$

Put

$$t := \frac{\kappa_1}{2}, \qquad m_i := \kappa_1 - \kappa_i \quad (1 \leq i \leq d-1), \qquad \mathfrak{m} := \sum_{i=2}^{d-1} m_i.$$

Thus $m_1 = 0$ and

$$0 = m_1 \leq m_2 \leq \cdots \leq m_{d-1}.$$

The choice $t = \kappa_1/2$ reflects the principal-ray normalization. In a determinant-bounded region, matrices of large norm are close to $\{\det = 0\}$. On compact singular-value windows one has, after separating the common scale,

$$\kappa_1(v), \ldots, \kappa_{d-1}(v) = 2t + O(1), \qquad \kappa_d(v) = -2(d-1)t + O(1).$$

The diagonal b_t therefore normalizes this configuration under the left–right action $a_t = b_t \otimes b_t$: the $d-1$ large singular directions are scaled down, while the small compensating direction is scaled up. Hence the principal ray supplies the only dynamical equidistribution input needed for the main term.

The remaining singular-value ratios are not treated by additional diagonal equidistribution. On compact windows they stay in a fixed compact parameter set, so all norm-comparison, weight, and Jacobian factors are uniformly bounded above and below. The regions where these ratios become unbalanced are controlled instead by the shell estimates below. We divide the cusp into two regimes: moderate degeneration, treated by a box with one long coordinate, and deep degeneration, treated in a chart centered at an off-diagonal Weyl element.

For $D > 0$, define

$$\mathcal{S}(\boldsymbol{\kappa}, D) := \left\{ v \in \mathrm{M}_d(\mathbb{R}) : \begin{array}{ll} \kappa_i - 1 < \kappa_i(v) \le \kappa_i & (1 \le i \le d-1), \\ |\det v| \le D \end{array} \right\}.$$

Every matrix in this shell has rank at least $d-1$.

For $R > 0$, set

$$\mathcal{B}_{\mathfrak{m},D}(R) := \left\{ x = (x_{ij}) \in \mathrm{M}_d(\mathbb{R}) : \begin{array}{ll} |x_{ij}| \le R, & (i,j) \ne (d,d), \\ |x_{dd}| \le R(1+D)e^{\mathfrak{m}}. \end{array} \right\}.$$

Lemma 12.4 (Moderate shells: a box with one long direction)**.** *For every $D > 0$, there exist constants $R = R_{d,D} > 1$ and $c = c_{d,D} > 0$ such that, for every $\boldsymbol{\kappa}$ as above and every $v \in \mathcal{S}(\boldsymbol{\kappa}, D)$,*

$$\mathrm{vol}_K\{k \in K : a_t k \cdot v \in \mathcal{B}_{\mathfrak{m},D}(R)\} \ge c \exp\left(-2 \sum_{i=1}^{d-1} \left(\frac{d}{2}\kappa_1 - m_i \right)_+ \right), \quad (12.3)$$

where $x_+ := \max\{x, 0\}$.

Proof. All implied constants depend only on d and D. By the singular-value decomposition and the left and right invariance of Haar measure on K, we may assume that

$$v = \mathrm{diag}\left(e^{\lambda_1}, \ldots, e^{\lambda_{d-1}}, \varepsilon e^{\lambda_d}\right),$$

where $\varepsilon \in \{-1, 0, 1\}$, $\kappa_i - 1 < \lambda_i \le \kappa_i$ for $1 \le i \le d-1$, and the final diagonal entry is understood to be zero when $\varepsilon = 0$.

Put

$$q_i := \frac{d}{2}\kappa_1 - m_i = (d-2)t + \kappa_i, \qquad \rho_i := \varepsilon_0 e^{-(q_i)_+},$$

where $\varepsilon_0 > 0$ is sufficiently small, depending only on d. For $x = (x_1, \ldots, x_{d-1})$, set

$$\mathsf{k}(x) := \sum_{i=1}^{d-1} x_i (E_{id} - E_{di}) \in \mathfrak{so}(d).$$

Choose a fixed sufficiently small neighborhood $\mathscr{U}$ of the identity in $\mathrm{SO}(d-1)$. In a fixed neighborhood of the identity in $\mathrm{SO}(d)$, the map

$$(h, x) \to \begin{pmatrix} h & 0 \\ 0 & 1 \end{pmatrix} \exp \mathsf{k}(x)$$

is a smooth coordinate chart with Jacobian bounded above and below. Let $\mathcal{K}_{\boldsymbol{\kappa}}$ be the image of

$$h \in \mathscr{U}, \qquad |x_i| \le \rho_i \quad (1 \le i \le d-1).$$

Then

$$\mathrm{vol}_{\mathrm{SO}(d)}(\mathcal{K}_{\boldsymbol{\kappa}}) \gg \prod_{i=1}^{d-1} \rho_i = \exp\left(-\sum_{i=1}^{d-1} (q_i)_+ \right). \quad (12.4)$$

We claim that

$$a_t(k_1, k_2) \cdot v \in \mathcal{B}_{\mathfrak{m},D}(R)$$

for all $k_1, k_2 \in \mathcal{K}_{\boldsymbol{\kappa}}$, provided that R is sufficiently large. Entries in the upper-left $(d-1)\times(d-1)$ block are $O(1)$, because $e^{-2t+\lambda_i} = O(1)$. For a mixed entry, the contribution of the i-th singular direction is

$$O\big(e^{(d-2)t+\lambda_i}\rho_i\big) = O(e^{q_i}\rho_i) = O(1).$$

The last singular direction satisfies the same bound because $\lambda_d \le \lambda_i$.

For the (d,d)-entry, the contribution of the last singular direction is

$$e^{2(d-1)t+\lambda_d} = |\det v| \exp\left((d-1)\kappa_1 - \sum_{i=1}^{d-1}\lambda_i\right) = O_D(e^{\mathfrak{m}}).$$

For $i<d$, the corresponding contribution is $O(e^{2(d-1)t+\lambda_i}\rho_i^2)$. If $q_i \ge 0$, this is

$$O\left(\exp(2(d-1)t + \kappa_i - 2q_i)\right) = O(e^{m_i}) = O(e^{\mathfrak{m}}).$$

If $q_i < 0$, then $\rho_i \asymp 1$, while $m_i > d\kappa_1/2$, and hence

$$2(d-1)t + \lambda_i \le d\kappa_1 - m_i < m_i \le \mathfrak{m}.$$

This proves the claim. Taking the product of (12.4) for the two factors of K yields (12.3). □

The preceding lemma is efficient only while $m_{d-1} \le d\kappa_1/2$. Beyond this threshold, the last two singular directions should be placed off the diagonal. Let $w \in \mathrm{SO}(d)$ be the identity on $\langle e_1, \dots, e_{d-2}\rangle$ and the rotation by $\pi/2$ on $\langle e_{d-1}, e_d\rangle$, so that

$$we_{d-1} = e_d, \qquad we_d = -e_{d-1}.$$

Lemma 12.5 (Deep shells: an off-diagonal Weyl chart)**.** *There exist constants $R_d > 1$ and $c_d > 0$ such that the following holds. Let $\tau \ge 1$, and let $v \in \mathrm{M}_d(\mathbb{R})$ satisfy*

$$\tau - 1 < \kappa_1(v) \le \tau, \qquad \kappa_{d-1}(v) \le -\frac{d-2}{2}\tau. \tag{12.5}$$

Then

$$\mathrm{vol}_K\left\{k \in K : \|a_{\tau/2}k\cdot v\| \le R_d\right\} \ge c_d e^{-\delta_d\tau}, \tag{12.6}$$

where $\delta_d := d(d-1) - \frac{d}{2}$. The convention $\kappa_{d-1}(v) = -\infty$ is allowed, so the lemma also covers matrices of rank at most $d-2$.

Proof. Write the singular values of v as $s_1 \ge \cdots \ge s_d \ge 0$. By the singular-value decomposition and Haar invariance, we may replace v by $D = \mathrm{diag}(s_1, \dots, s_d)$, with the sign of the final entry changed if necessary. Put $\rho := \varepsilon_0 e^{-d\tau/2}$, where $\varepsilon_0 > 0$ is a sufficiently small dimensional constant.

For $1 \le i \le d-1$, write $X_i = E_{id} - E_{di}$. Near the identity, and near w, use the coordinate charts

$$k_1 = \begin{pmatrix} h_1 & 0 \\ 0 & 1\end{pmatrix}\exp\left(\sum_{i=1}^{d-2} x_i X_i\right)\exp(\alpha X_{d-1}),$$

$$k_2 = \begin{pmatrix} h_2 & 0 \\ 0 & 1\end{pmatrix}\exp\left(\sum_{i=1}^{d-2} y_i X_i\right)\exp(\beta X_{d-1})w,$$

where h_1, h_2 range over fixed small neighborhoods of the identity in $\mathrm{SO}(d-1)$. These are full-dimensional smooth charts, and their Jacobians are bounded above and below on fixed coordinate boxes.

Restrict

$$|x_i|, |y_i| \le \rho \qquad (1 \le i \le d-2),$$

and $|\alpha|, |\beta| \le \varepsilon_0$. In the last two singular directions, the (d,d)-entry before applying the diagonal weights is

$$F(\alpha, \beta) := s_{d-1} \sin\alpha \cos\beta - s_d \cos\alpha \sin\beta.$$

We impose the additional condition

$$|F(\alpha, \beta)| \le \varepsilon_0 e^{-(d-1)\tau}. \tag{12.7}$$

If $s_{d-1} = 0$, this condition is automatic. Otherwise, after reducing ε_0, one has $|\partial_\alpha F(\alpha, \beta)| \asymp s_{d-1}$ throughout the coordinate square. Fubini's theorem and the one-dimensional mean-value theorem therefore show that the area of the set of (α, β) satisfying (12.7) is

$$\gg \min\left\{1, \frac{e^{-(d-1)\tau}}{s_{d-1}}\right\}.$$

By (12.5), $s_{d-1} \le e^{-(d-2)\tau/2}$, and hence this area is $\gg e^{-d\tau/2}$. It follows that the set of pairs (k_1, k_2) just described has Haar measure

$$\gg \rho^{2(d-2)} e^{-d\tau/2} = \exp\left(-\left(d(d-2) + \frac{d}{2}\right)\tau\right) = e^{-\delta_d \tau}. \tag{12.8}$$

It remains to bound the image of D. An entry in the upper-left $(d-1) \times (d-1)$ block of $b_t k_1 D k_2^{\intercal} b_t$ is bounded by $O(e^{-\tau} s_1) = O(1)$. A mixed entry has weight $e^{(d-2)\tau/2}$. The contribution of a singular direction $i \le d-2$ contains one of the factors x_i, y_i, and is therefore

$$O\left(e^{(d-2)\tau/2} s_i \rho\right) = O(1).$$

The contributions of the last two directions are also $O(1)$, by the bound on s_{d-1}.

The (d,d)-entry has weight $e^{(d-1)\tau}$. Its contribution from the last two directions is bounded by (12.7). For $i \le d-2$, both a left and a right small coordinate occur, and the contribution is

$$O\left(e^{(d-1)\tau} s_i \rho^2\right) = O\left(e^{d\tau} e^{-d\tau}\right) = O(1).$$

The remaining Taylor terms contain either the same two small coordinates, or one small coordinate together with s_{d-1}; the latter are bounded by

$$O\left(e^{(d-1)\tau} s_{d-1} \rho\right) = O(1).$$

Thus all entries are uniformly bounded on the coordinate set. Enlarging a dimensional constant gives $\|a_t(k_1, k_2) \cdot D\| \le R_d$. Together with (12.8), this proves (12.6). $\square$

We now turn the two lower bounds into lattice-point estimates.

Lemma 12.6 (Counting moderate and deep shells)**.** *Let $D > 0$. The following estimates hold uniformly in their stated parameters.*

(i) *If $m_{d-1} \le \frac{d}{2}\kappa_1$, then*

$$\#\big(\Lambda_{\mathrm{ni}} \cap \mathcal{S}(\boldsymbol{\kappa}, D)\big) \ll_{\Lambda,D} \exp\big(d(d-1)\kappa_1 - \mathfrak{m}\big). \tag{12.9}$$

(ii) *If $\tau \ge 1$, then*

$$\#\left\{v \in \Lambda_{\mathrm{ni}} : \begin{array}{l} \tau - 1 < \kappa_1(v) \le \tau, \\ \kappa_{d-1}(v) \le -(d-2)\tau/2 \end{array}\right\} \ll_\Lambda e^{-d\tau/2} e^{d(d-1)\tau}. \tag{12.10}$$

Proof. Let $R = R_{d,D}$ be as in Lemma 12.4. By Lemma 12.1(ii), uniformly in $k \in K$,

$$\#\left\{v \in \Lambda_{\mathrm{ni}} : a_t k \cdot v \in \mathcal{B}_{\mathfrak{m},D}(R)\right\} \ll_D e^{\mathfrak{m}}\big(1 + \widehat{\alpha}_{\eta,M}(a_t k; \Lambda)\big).$$

Summing (12.3) over the vectors in the shell and integrating over K, we obtain

$$\#\big(\Lambda_{\mathrm{ni}} \cap \mathcal{S}(\boldsymbol{\kappa}, D)\big) \ll_D \exp\left(\mathfrak{m} + 2\sum_{i=1}^{d-1}\left(\frac{d}{2}\kappa_1 - m_i\right)_+\right) \int_K \big(1 + \widehat{\alpha}_{\eta,M}(a_t k; \Lambda)\big)\, dk.$$

The last integral is uniformly bounded by Theorem 3.13. Under the hypothesis in (i), every $m_i \le d\kappa_1/2$, and hence

$$2\sum_{i=1}^{d-1}\left(\frac{d}{2}\kappa_1 - m_i\right) = d(d-1)\kappa_1 - 2\mathfrak{m}.$$

This proves (12.9).

For (ii), choose a nonnegative $f \in C_c(\mathrm{M}_d(\mathbb{R}))$ which is identically one on the ball of radius R_d from Lemma 12.5. The lower bound (12.6) gives

$$1 \ll e^{\delta_d \tau} \int_K f(a_{\tau/2} k \cdot v)\, dk$$

for every vector counted in (12.10). Summing over v, applying Lemma 12.1(i), and then using Theorem 3.13, we obtain

$$\#\left\{v \in \Lambda_{\mathrm{ni}} : a_t k \cdot v \in \mathcal{B}_{\mathfrak{m},D}(R)\right\} \ll_\Lambda e^{\delta_d \tau} = e^{-d\tau/2} e^{d(d-1)\tau},$$

as required. □

The shell estimates imply both the compact-window bound needed in the kernel comparison and the cusp-tail bound needed for norm balls.

Corollary 12.7 (Counting in a compact singular-value window)**.** *Let $\mathscr{C} \subset B_0^+$ be compact and let $D > 0$. Then*

$$\#\left\{v \in \Lambda_{\mathrm{ni}} : T^{-1} b^0(v) \in \mathscr{C},\ |\det v| \le D\right\} \ll_{\Lambda,\mathscr{C},D} T^{d(d-1)}$$

uniformly for $T \ge 2$.

Proof. Compactness of $\mathscr{C}$ implies

$$\kappa_i(v) = \log T + O_{\mathscr{C}}(1) \qquad (1 \le i \le d-1).$$

Thus the relevant matrices lie in finitely many unit shells for which $\mathfrak{m} = O_{\mathscr{C}}(1)$. For all sufficiently large T, these are moderate shells. Applying (12.9) and summing over the finitely many shells gives the assertion. The bounded range of T is absorbed into the implied constant. □

Lemma 12.8 (An elementary deficit sum)**.** *For every $n \ge 2$, there exist a constant $C_n > 0$ such that*

$$\sum_{\substack{0=m_1\le m_2\le\cdots\le m_n\\ m_n\ge M}} \exp\left(-\sum_{i=2}^{n} m_i\right) \le C_n e^{-M/2} \qquad (M \ge 0).$$

Proof. For a fixed value $m_n = m$, the number of nondecreasing tuples $(m_2, \dots, m_{n-1})$ is at most $(m+1)^{n-2}$, while $\sum_{i=2}^{n} m_i \ge m$. Therefore the left-hand side is bounded by

$$\sum_{m\ge M} (m+1)^{n-2} e^{-m} \ll_n e^{-M/2}.$$

□

Proposition 12.9 (Uniform singular-value cusp tail)**.** *Fix $D > 0$ and a K-invariant norm on $\mathrm{M}_d(\mathbb{R})$. There exist constants $c, C > 0$, depending on Λ, D, and the norm, such that for every $R \ge 1$ and $T \ge 2$,*

$$\begin{aligned}\#\{v \in \Lambda_{\mathrm{ni}} : \|v\| < T,\ |\det v| \le D,\ \kappa_{d-1}(v) < \log T - R\}\\ \le Ce^{-R/2}T^{d(d-1)} + CT^{d(d-1)-d/2}. \quad (12.11)\end{aligned}$$

In particular,

$$\limsup_{T\to\infty} T^{-d(d-1)}\#\{\cdots\} \ll e^{-R/2}.$$

Proof. By equivalence of norms, there is $c_0 \ge 0$ such that

$$\kappa_1(v) \le \log\|v\| + c_0$$

for every nonzero v. Matrices with $\kappa_1(v) < 0$ lie in a fixed compact set and contribute $O_\Lambda(1)$. For the remaining matrices, let $\tau = \lceil\kappa_1(v)\rceil$. Then $1 \le \tau \le \log T + c_0 + 1$.

First consider the deep region $\kappa_{d-1}(v) \le -\frac{d-2}{2}\tau$. By (12.10), its contribution is

$$\ll_\Lambda \sum_{1\le\tau\le\log T+c_0+1} e^{(d(d-1)-d/2)\tau} \ll T^{d(d-1)-d/2}.$$

It remains to consider the moderate region. For $1 \le i \le d-1$, put

$$\tau_i := \lceil\kappa_i(v)\rceil, \qquad m_i := \tau - \tau_i.$$

Then

$$0 = m_1 \le m_2 \le \cdots \le m_{d-1} < \frac{d}{2}\tau,$$

and v belongs to the unit shell with upper endpoints $(\tau_1, \dots, \tau_n)$. By (12.9), this shell contains

$$\ll_{\Lambda,D} \exp\left(d(d-1)\tau - \sum_{i=2}^{n} m_i\right)$$

nonisotropic lattice points.

The cusp condition implies

$$m_{d-1} = \tau - \tau_{d-1} \geq \tau - \log T + R - 1 \geq R - \nu - C_0$$

for a constant C_0 depending only on the norm. Therefore Lemma 12.8 gives, for each fixed τ,

$$\sum_{\text{admissible } (m_2,\dots,m_{d-1})} \exp\left(-\sum_{i=2}^{d-1} m_i\right) \ll \exp\bigl(-\tfrac{1}{2}(R-\nu-C_0)_+\bigr).$$

Since $e^{d(d-1)\tau} \ll T^{d(d-1)} e^{-d(d-1)\nu}$, summing over $\nu \geq 0$ yields

$$\sum_{\nu \geq 0} T^{d(d-1)} e^{-d(d-1)\nu} \exp\bigl(-\tfrac{1}{2}(R-\nu-C_0)_+\bigr) \ll T^{d(d-1)} e^{-R/2}.$$

Indeed, for $\nu \leq R - C_0$ the summand is bounded by $e^{-R/2} e^{-(d(d-1)-c)\nu}$, while for $\nu > R - C_0$ the geometric factor $e^{-d(d-1)\nu}$ already gives exponential decay in R. Combining the moderate and deep regions proves (12.11). ☐

12.3. **Equidistribution of modified Siegel transforms.** For $k \in K$, put

$$\omega(k) := (k^{\mathsf{T}} \cdot E_{dd})^{\mathsf{T}} \in \mathbb{S}^{N-1}.$$

We use the following form of Shah's theorem.

Theorem 12.10 (Shah)**.** *Let $F \in C_c(X)$ and $\nu \in C(\mathbb{S}^{N-1})$. If $H\Lambda$ is dense in X, then*

$$\lim_{t\to\infty} \int_K F(a_t k\Lambda)\nu(\omega(k))\,dk = \left(\int_X F\,dm_X\right)\left(\int_K \nu(\omega(k))\,dk\right).$$

Proof. This is [30, Corollary 1.2]; see also [29] for the measure-classification input. ☐

We now extend the theorem to the modified Siegel transform. The proof is a direct application of Lemma 12.2; this avoids a separate lower- and upper-bound argument.

Theorem 12.11 (Equidistribution for modified Siegel transforms)**.** *Let $f \in C_c(\mathsf{M}_d^+)$, and let $\nu \in C(\mathbb{S}^{N-1})$. Suppose that Λ is Diophantine and not determinant-rational. Then*

$$\lim_{t\to\infty} \int_K \widetilde{f}_{\mathrm{ni}}(a_t k;\Lambda)\nu(\omega(k))\,dk = \left(\int_{\mathrm{M}_d(\mathbb{R})} f(v)\,dv\right)\left(\int_K \nu(\omega(k))\,dk\right). \tag{12.12}$$

Proof. By linearity, it is enough to treat nonnegative f and nonnegative ν. Since Λ is not determinant-rational, Theorem 2.9 implies that $H\Lambda$ is dense in X.

For $t \geq 0$, define the finite measure μ_t on X by

$$\int_X F\, d\mu_t := \int_K F(a_t k\Lambda)\nu(\omega(k))\, dk.$$

By Theorem 12.10,

$$\mu_t \to \left(\int_K \nu(\omega(k))\, dk\right) m_X \quad \text{weakly as } t \to \infty.$$

Put $\Phi = \widetilde{f}_{\mathrm{ni}}$. By Lemma 12.3, Φ is Borel measurable and is continuous outside an m_X-null set. By Lemma 12.1(i),

$$|\Phi(a_t k\Lambda)| \ll_f 1 + \widehat{\alpha}_{\eta,M}(a_t k; \Lambda).$$

Choose $\theta > 0$ as in Theorem 3.13. Since ν is bounded, that theorem gives

$$\sup_{t \geq 0} \int_X |\Phi|^{1+\theta}\, d\mu_t < \infty.$$

The bounded range of t follows from the log-Lipschitz property of the height. Thus Lemma 12.2 applies and gives

$$\lim_{t\to\infty} \int_X \Phi\, d\mu_t = \left(\int_K \nu(\omega(k))\, dk\right) \int_X \Phi\, dm_X.$$

Finally, Lemma 12.3 and Siegel's formula imply

$$\int_X \Phi\, dm_X = \int_X \widetilde{f}\, dm_X = \int_{\mathrm{M}_d(\mathbb{R})} f(v)\, dv.$$

This proves (12.12). $\square$

12.4. **Comparison with the fiber kernel.** We next combine the pointwise comparison from Proposition 11.3 with the compact singular-value window count proved above.

Proposition 12.12. *Let $f \in C_c(\mathsf{M}_d^+)$, and let $\nu \in C(\mathbb{S}^{N-1})$. Then as $T \to \infty$,*

$$T^{-d(d-1)} \sum_{\substack{v\in\Lambda_{\mathrm{ni}} \\ \widehat{v}\neq 0}} J_f(T^{-1}b^0(v), \det v)\nu\left(\frac{\widehat{v}}{\|\widehat{v}\|}\right)$$
$$- \omega_{d-1}^2 \int_K \widetilde{f}_{\mathrm{ni}}(a_{\frac{1}{2}\log T}k; \Lambda)\nu(\omega(k))\, dk \to 0. \qquad (12.13)$$

Proof. Put $t = \frac{1}{2}\log T$. Let $\Xi_T \subset \Lambda_{\mathrm{ni}}$ be the set of all v for which either

$$J_f(T^{-1}b^0(v), \det v) \neq 0 \text{ or } f(a_t k \cdot v) \neq 0 \quad \text{for some } k \in K.$$

Compactness of the supports in Proposition 11.3 gives a compact set $\mathscr{C}_f \subset B_0^+$ and a constant $D_f > 0$ such that

$$v \in \Xi_T \quad \Longrightarrow \quad T^{-1}b^0(v) \in \mathscr{C}_f, \qquad |\det v| \leq D_f.$$

Therefore Corollary 12.7 gives

$$\#\Xi_T \ll_f T^{d(d-1)}. \tag{12.14}$$

Moreover, every $v \in \Xi_T$ has rank at least $d-1$, and hence $\widehat{v} \neq 0$. The same support localization gives $\|v\| \gg_f T$.

Fix $\varepsilon > 0$. Apply Proposition 11.3 with an error $\varepsilon_1 > 0$ to be chosen. For all sufficiently large T, the resulting estimate is valid for every $v \in \Xi_T$, and gives

$$\left| \omega_{d-1}^2 T^{d(d-1)} \int_K f(a_t k \cdot v)\nu(\omega(k))\, dk - J_f(T^{-1}b^0(v), \det v)\nu\left(\frac{\widehat{v}}{\|\widehat{v}\|}\right) \right| < \varepsilon_1.$$

Sum this inequality over $v \in \Xi_T$, multiply by $T^{-d(d-1)}$, and use (12.14). Choosing ε_1 so that the resulting upper bound is less than ε proves (12.13). □

Proposition 12.13 (Counting in the $(b^0, \det)$-parameter space)**.** *Suppose that Λ is Diophantine and non-determinant-rational. Let $h \in C_c(B_0^+ \times \mathbb{R})$ and $\nu \in C(\mathbb{S}^{N-1})$. Then*

$$\lim_{T\to\infty} T^{-d(d-1)} \sum_{\substack{v\in\Lambda_{\mathrm{ni}} \\ \widehat{v}\neq 0}} h(T^{-1}b^0(v), \det v)\nu\left(\frac{\widehat{v}}{\|\widehat{v}\|}\right)$$

$$= \omega_{d-1}^2 \left(\int_K \nu(\omega(k))\, dk\right)\left(\int_{\mathbb{R}}\int_{B_0^+} h(r,\zeta)\det(r)^d\, dr\, d\zeta\right). \tag{12.15}$$

Proof. By linearity and approximation, it suffices to treat nonnegative h and ν. By Lemma 11.2, choose $f \in C_c(\mathsf{M}_d^+)$ such that $J_f = h$. Applying Proposition 12.12, followed by Theorem 12.11, gives the right-hand side of (12.15) with the final factor $\int_{\mathrm{M}_d(\mathbb{R})} f$. By Lemma 11.5,

$$\int_{\mathrm{M}_d(\mathbb{R})} f(v)\, dv = \int_{\mathbb{R}}\int_{B_0^+} h(r,\zeta)\det(r)^d\, dr\, d\zeta.$$

□

Corollary 12.14. *Let $\varphi \in C_c(B_0^+)$ and $\psi \in C_c(\mathbb{R})$. Then*

$$\lim_{T\to\infty} T^{-d(d-1)} \sum_{\substack{v\in\Lambda_{\mathrm{ni}} \\ \widehat{v}\neq 0}} \varphi(T^{-1}b^0(v))\psi(\det v)$$

$$= \omega_{d-1}^2 \left(\int_{B_0^+} \varphi(r)\det(r)^d\, dr\right)\left(\int_{\mathbb{R}} \psi(\zeta)\, d\zeta\right).$$

Proof. Apply Proposition 12.13 with $h(r,\zeta) = \varphi(r)\psi(\zeta)$ and $\nu \equiv 1$. □

The same conclusion holds for bounded compactly supported Borel functions whose discontinuity set has measure zero with respect to

$$\det(r)^d\, dr\, d\zeta.$$

Indeed, approximate such a function from above and below by continuous compactly supported functions and use Proposition 12.13. We shall use this extension without further comment.

12.5. **Passage from compact windows to norm balls.** Fix a K-invariant norm $\|\cdot\|$ on $\mathrm{M}_d(\mathbb{R})$. For $r \in B_0^+$ and $\tau \in \mathbb{R}$, put

$$\Phi(r,\tau) := \left\|\mathrm{diag}\left(r, \frac{\tau}{\det r}\right)\right\|.$$

Because the norm is K-invariant, for every matrix v of rank at least $d-1$,

$$\|v\| = \Phi(b^0(v), \det v). \tag{12.16}$$

Consequently,

$$\|v\| < T \quad \Longleftrightarrow \quad \Phi(T^{-1}b^0(v), T^{-d}\det v) < 1. \tag{12.17}$$

Theorem 12.15 (Counting in norm balls)**.** *Suppose that $\Lambda \in X$ is Diophantine and non-determinant-rational. Then, for every $a < b$,*

$$\lim_{T\to\infty} T^{-d(d-1)} \#\{v \in \Lambda_{\mathrm{ni}} : \|v\| < T,\ a < \det v < b\} = C_{\|\cdot\|}(b-a), \tag{12.18}$$

where

$$C_{\|\cdot\|} := \omega_{d-1}^2 \int_{\{r\in B_0^+:\ \|\mathrm{diag}(r,0)\|<1\}} \det(r)^d\, dr. \tag{12.19}$$

Proof. Fix $R > 1$, and put

$$D_R := \left\{r \in B_0^+ : r_{d-1} \geq e^{-R},\ \Phi(r,0) < 1\right\}.$$

This is relatively compact in B_0^+. Define

$$N_R(T) := \#\left\{v \in \Lambda_{\mathrm{ni}} : \begin{array}{l} \|v\| < T,\ a < \det v < b, \\ \kappa_{d-1}(v) \geq \log T - R \end{array}\right\}.$$

For fixed R, the convergence

$$\Phi(r, T^{-d}\zeta) \to \Phi(r,0)$$

is uniform for r in compact subsets with $r_{d-1} \geq e^{-R}$ and $\zeta \in [a,b]$. Hence (12.17), the boundary-null extension of Corollary 12.14, and a standard inner–outer approximation give

$$\lim_{T\to\infty} T^{-d(d-1)} N_R(T) = \omega_{d-1}^2 (b-a) \int_{D_R} \det(r)^d\, dr. \tag{12.20}$$

Here the relevant boundary is null because it is contained in the union of $\{r_{d-1} = e^{-R}\}$ and the boundary of the unit ball of the norm restricted to the diagonal subspace.

Let

$$E_R(T) := \#\left\{v \in \Lambda_{\mathrm{ni}} : \begin{array}{l} \|v\| < T,\ a < \det v < b, \\ \kappa_{d-1}(v) < \log T - R \end{array}\right\}.$$

Apply Proposition 12.9 with $D = 1 + \max\{|a|,|b|\}$. It gives

$$\limsup_{T\to\infty} T^{-d(d-1)} E_R(T) \ll e^{-R/2}. \tag{12.21}$$

Since the full counting function is $N_R(T)+E_R(T)$, (12.20) and (12.21) imply

$$\omega_{d-1}^2(b-a)\int_{D_R}\det(r)^d\,dr \le \liminf_{T\to\infty} T^{-d(d-1)}N_T(a,b)$$

and

$$\limsup_{T\to\infty} T^{-d(d-1)}N_\Lambda(a,b;T) \le \omega_{d-1}^2(b-a)\int_{D_R}\det(r)^d\,dr + O(e^{-R/2}),$$

where

$$N_\Lambda(a,b;T) := \#\{v\in\Lambda_{\mathrm{ni}} : \|v\|<T,\ a<\det v<b\}.$$

As $R\to\infty$, the sets D_R increase to

$$\{r\in B_0^+ : \Phi(r,0)<1\}.$$

Monotone convergence and then $R\to\infty$ prove (12.18). □

Lemma 12.16 (The corresponding volume asymptotic)**.** *For every $a<b$,*

$$\mathrm{vol}\{v\in \mathrm{M}_d(\mathbb{R}) : \|v\|<T,\ a<\det v<b\} \sim C_{\|\cdot\|}(b-a)T^{d(d-1)}. \tag{12.22}$$

Proof. The rank-at-most-$(d-2)$ locus has Lebesgue measure zero. On its complement, use (12.17) and the integral identity Lemma 11.6. For each fixed R, the same compact-window approximation used in (12.20) gives

$$\lim_{T\to\infty} T^{-d(d-1)}V_R(T) = \omega_{d-1}^2(b-a)\int_{D_R}\det(r)^d\,dr,$$

where $V_R(T)$ denotes the volume with the additional condition $\kappa_{d-1}(v)\ge \log T - R$.

The omitted part corresponds in the limiting integral to

$$\left\{r\in B_0^+ : \Phi(r,0)<1,\ r_{d-1}<e^{-R}\right\}.$$

Its contribution tends to zero as $R\to\infty$, by dominated convergence; indeed the region is bounded and the density is $\det(r)^d$. Letting $R\to\infty$ yields (12.22). □

12.6. **Completion of the counting theorems.** We finish by assembling the regular and singular contributions. Put

$$S_{\mathrm{ni}}(T) := \#\{v\in\Lambda_{\mathrm{ni}} : \|v\|<T,\ \det v = 0\}.$$

For every $\varepsilon>0$,

$$S_{\mathrm{ni}}(T) \le \#\{v\in\Lambda_{\mathrm{ni}} : \|v\|<T,\ -\varepsilon<\det v<\varepsilon\}.$$

By Theorem 12.15,

$$\limsup_{T\to\infty} T^{-d(d-1)}S_{\mathrm{ni}}(T) \le 2C_{\|\cdot\|}\varepsilon.$$

Letting $\varepsilon\to 0$, we obtain

$$S_{\mathrm{ni}}(T) = o(T^{d(d-1)}). \tag{12.23}$$

Every lattice point with nonzero determinant belongs to Λ_{ni}. Hence, for unimodular Λ,

$$\#\{v \in \Lambda : \|v\| < T,\ a < \det v < b,\ \det v \neq 0\} = C_{\|\cdot\|}(b-a)T^{d(d-1)} + o(T^{d(d-1)}).$$

For a general lattice, put $c = \mathrm{covol}(\Lambda)^{-1/d^2}$, so that $c\Lambda$ is unimodular. Since

$$\det(cv) = c^d \det v, \qquad \|cv\| = c\|v\|, \qquad c^{d+d(d-1)} = c^{d^2} = \mathrm{covol}(\Lambda)^{-1},$$

we obtain

$$\begin{aligned} &\#\{v \in \Lambda : \|v\| < T,\ a < \det v < b,\ \det v \neq 0\} \\ &= \frac{C_{\|\cdot\|}}{\mathrm{covol}(\Lambda)}(b-a)T^{d(d-1)} + o(T^{d(d-1)}). \end{aligned} \tag{12.24}$$

Together with Lemma 12.16, this proves Theorem 3.8 and the volume formulation in Theorem 1.1. The determinant-form statement follows from the coordinate change defining $F_{\Lambda,\mathcal{B}}$, as recorded in Theorem 3.10.

Assume in addition that Λ satisfies the isotropic noncoincidence condition. Since a non-determinant-rational lattice is not of $\mathbb{Q}$-split type, Theorem 3.15 gives

$$\#\{v \in \Lambda_{\mathrm{iso}} : \|v\| < T\} = c_\Lambda^{\mathrm{sing}} T^{d(d-1)} + o(T^{d(d-1)}).$$

This together with (12.23) proves (1.4). The finite-piece decomposition in Lemma 10.7, together with the positive main constants in Lemma 10.6, gives the stated positivity criterion for $c_\Lambda^{\mathrm{sing}}$. Finally, adding the regular and singular contributions proves (1.5) and hence Theorem 1.2.

13. Algebraic lattices are Diophantine

In this section we prove Proposition 3.7 and record several examples. The proof is an application of a Liouville estimate to the integral Plücker vectors associated with rational subspaces of an algebraic lattice.

Lemma 13.1 (Liouville estimate for algebraic linear maps)**.** *Let*

$$A : \mathbb{R}^s \to \mathbb{R}^t$$

be a linear map whose matrix entries are algebraic numbers. Then there are constants $c > 0$ and $M \geq 0$, depending only on A, such that

$$Az \neq 0 \quad \Longrightarrow \quad \|Az\| \geq c \max\{1, \|z\|\}^{-M} \qquad (z \in \mathbb{Z}^s).$$

Proof. Choose a number field $K \subset \mathbb{R}$ containing the entries of A, and let $\sigma_0 : K \hookrightarrow \mathbb{R}$ be the given embedding. Choose an integer $D \geq 1$ such that every entry of DA is an algebraic integer.

Suppose that $Az \neq 0$. Some coordinate α of Az is then nonzero, and $D\alpha$ is a nonzero algebraic integer. Hence

$$1 \leq \left|N_{K/\mathbb{Q}}(D\alpha)\right| = \prod_{\sigma : K \hookrightarrow \mathbb{C}} \left|\sigma(D\alpha)\right|.$$

For every embedding $\sigma : K \hookrightarrow \mathbb{C}$, the entries of $\sigma(DA)$ are bounded by a constant depending only on A. Thus

$$\bigl|\sigma(D\alpha)\bigr| \le C \max\{1, \|z\|\}$$

for a constant $C \ge 1$ independent of z. Isolating the factor corresponding to σ_0, we obtain

$$|D\alpha| \ge C^{1-[K:\mathbb{Q}]} \max\{1, \|z\|\}^{1-[K:\mathbb{Q}]}.$$

Since $\|Az\| \ge |\alpha|$, the assertion follows. □

Proof of Proposition 3.7. As before, identify $\mathrm{M}_d(\mathbb{R})$ with $\mathbb{R}^N$. Choose a $\mathbb{Z}$-basis of Λ and write

$$\Lambda = g\mathbb{Z}^N, \qquad g \in \mathrm{GL}_N(\overline{\mathbb{Q}} \cap \mathbb{R}).$$

We first note that the projections appearing in Definition 3.6 are defined over $\mathbb{Q}$. Indeed, the skew Cauchy decomposition in (3.3) is obtained by base change from the corresponding decomposition over $\mathbb{Q}$, and each of its Schur-functor summands is defined over $\mathbb{Q}$; see, for example, [1, 15]. Consequently,

$$\pi_{k,m} : \wedge^{kd}\mathbb{R}^N \to \mathcal{M}_{k,m}$$

and $\pi_{k,m}^{(r)} = \wedge^r \pi_{k,m}$ have rational matrices with respect to the standard exterior bases.

Fix $1 \le k \le d-1$, $m \in \{1, 2\}$, and $r \in \{1, \dim \mathcal{M}_{k,m}\}$. Let $V_1, \dots, V_r$ be kd-dimensional Λ-rational subspaces. For each j, set

$$L_j := g^{-1}(\Lambda \cap V_j) = \mathbb{Z}^N \cap g^{-1}V_j.$$

Then L_j is a primitive sublattice of $\mathbb{Z}^N$ of rank kd. Choose a $\mathbb{Z}$-basis of L_j, and let $q_j \in \wedge^{kd}\mathbb{Z}^N$ be the corresponding primitive Plücker vector, with its sign chosen so that

$$\mathsf{w}_{\Lambda,V_j} = (\wedge^{kd} g) q_j.$$

Define

$$q := q_1 \wedge \cdots \wedge q_r \in \wedge^r \left(\wedge^{kd}\mathbb{Z}^N\right) \quad \text{and} \quad \mathbf{w} := \mathsf{w}_{\Lambda,V_1} \wedge \cdots \wedge \mathsf{w}_{\Lambda,V_r}.$$

Then

$$\mathbf{w} = B_{k,r} q, \qquad B_{k,r} := \wedge^r \left(\wedge^{kd} g\right). \tag{13.1}$$

If $q = 0$, then $\mathbf{w} = 0$, so the first alternative in Definition 3.6 holds. Suppose henceforth that $q \ne 0$. Since $B_{k,r}$ is invertible, there is a constant $C_{k,r} \ge 1$ such that

$$\|q\| \le C_{k,r} \|\mathbf{w}\|. \tag{13.2}$$

In particular, because q is a nonzero integral vector,

$$\|\mathbf{w}\| \ge C_{k,r}^{-1}. \tag{13.3}$$

Consider the fixed linear map

$$A_{k,m,r} := \bigl(I - \pi_{k,m}^{(r)}\bigr) B_{k,m}.$$

Its matrix entries are algebraic, and by (13.1),

$$\mathbf{w} - \pi_{k,m}^{(r)}(\mathbf{w}) = A_{k,m,r} q.$$

If this vector is nonzero, Lemma 13.1 gives constants $c_{k,m,r} > 0$ and $M_{k,m,r} \geq 0$ such that

$$\begin{aligned}\|\mathbf{w} - \pi_{k,m}^{(r)}(\mathbf{w})\| &\geq c_{k,m,r} \|q\|^{-M_{k,m,r}} \\ &\geq c_{k,m,r} C_{k,r}^{-M_{k,m,r}} \|\mathbf{w}\|^{-M_{k,m,r}}.\end{aligned}$$

There are only finitely many triples (k, m, r). Let $c_0 > 0$ be a common lower bound for the nonzero vectors $\mathbf{w}$ furnished by (13.3), and choose

$$M_0 > \max\big\{1, M_{k,m,r} : 1 \leq k \leq d-1,\ m \in \{1,2\},\ r \in \{1, q_{k,m}\}\big\}.$$

For each admissible triple,

$$\|\mathbf{w}\|^{-M_{k,m,r}} = \|\mathbf{w}\|^{-M_0} \|\mathbf{w}\|^{M_0 - M_{k,m,r}} \geq \min\{1, c_0^{M_0 - M_{k,m,r}}\} \|\mathbf{w}\|^{-M_0}.$$

Taking the minimum of the resulting finitely many positive constants gives $\eta_0 > 0$ such that

$$\|\mathbf{w} - \pi_{k,m}^{(r)}(\mathbf{w})\| \geq \eta_0 \|\mathbf{w}\|^{-M_0}$$

whenever $\mathbf{w} \neq \pi_{k,m}^{(r)}(\mathbf{w})$. This is precisely the condition in Definition 3.6. □

Corollary 13.2. *Every determinant form with algebraic coefficients is Diophantine.*

Proof. Let F be such a form. By definition, there exists $g \in \mathrm{GL}_N(\mathbb{R})$ such that $F = \det \circ g$. The polynomial identity $F = \det \circ g$, together with the condition $g \in \mathrm{GL}_N$, is an algebraic condition on the entries of g with coefficients in $\overline{\mathbb{Q}} \cap \mathbb{R}$. More explicitly, after comparing coefficients and adding an auxiliary variable to express $\det g \neq 0$, this condition becomes an existential first-order formula over the real closed field $\overline{\mathbb{Q}} \cap \mathbb{R}$. Since it has a solution over $\mathbb{R}$, Tarski's transfer principle for real closed fields [4] gives a solution

$$g_0 \in \mathrm{GL}_N(\overline{\mathbb{Q}} \cap \mathbb{R}).$$

Thus F is realized by the algebraic lattice $g_0 \mathbb{Z}^N$. The conclusion follows from Proposition 3.7 and Definition 3.9. □

Consequently, Theorem 1.1 follows from Theorem 3.8 and Proposition 3.7; the determinant-form version follows from Theorem 3.10 and Corollary 13.2.

13.1. **Examples.**

Example 13.3 (Entrywise algebraic lattices)**.** Let $\alpha, \beta, \gamma, \delta \in \overline{\mathbb{Q}} \cap \mathbb{R}$ be nonzero, and define the lattice

$$\Lambda_{\alpha,\beta,\gamma,\delta} := \begin{pmatrix} \alpha\mathbb{Z} & \beta\mathbb{Z} & \mathbb{Z} \\ \gamma\mathbb{Z} & \delta\mathbb{Z} & \mathbb{Z} \\ \mathbb{Z} & \mathbb{Z} & \mathbb{Z} \end{pmatrix} < \mathrm{M}_3(\mathbb{R}).$$

Unless all $\alpha, \beta, \gamma, \delta$ are rational, $\Lambda_{\alpha,\beta,\gamma,\delta}$ is not determinant rational and hence Theorem 1.1 applies. To see this, note that, with respect to the evident basis, the six coefficients of the determinant corresponding to the six permutations are, up to sign,

$$\alpha\delta, \quad \alpha, \quad \beta\gamma, \quad \beta, \quad \gamma, \quad \delta.$$

If the lattice were determinant-rational, these six numbers would be proportional over $\mathbb{Q}$. Taking the ratios $(\alpha\delta)/\alpha = \delta$ and $(\beta\gamma)/\beta = \gamma$, and then comparing α with δ and β with γ, shows that $\alpha, \beta, \gamma, \delta$ would all be rational.

The rank-six submodule obtained by setting the first row equal to zero is contained in a column-isotropic subspace. Since $d = 3$, the isotropic noncoincidence condition is vacuous. Thus Theorem 3.15 applies, and the usual lattice-point asymptotic in this rank-six submodule gives

$$c^{\mathrm{sing}}_{\Lambda_{\alpha,\beta,\gamma,\delta}} > 0.$$

Example 13.4 (A cubic norm-form lattice)**.** Let $K = \mathbb{Q}(\theta)$, where $\theta^3 = 2$, and let

$$\iota : K \hookrightarrow \mathrm{End}_{\mathbb{Q}}(K) \simeq \mathrm{M}_3(\mathbb{Q})$$

be the regular representation with respect to the basis $1, \theta, \theta^2$. Then $\iota(1) = I$ and

$$\iota(\theta) = T := \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \iota(\theta^2) = T^2 = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ 1 & 0 & 0 \end{pmatrix}.$$

In particular,

$$\det(xI + yT + zT^2) = N_{K/\mathbb{Q}}(x + y\theta + z\theta^2) = x^3 + 2y^3 + 4z^3 - 6xyz.$$

Let $\alpha \in \overline{\mathbb{Q}} \cap \mathbb{R}$ satisfy $\alpha^3 \notin \mathbb{Q}$, and put

$$W := \{A \in \mathrm{M}_3(\mathbb{Q}) : \text{the first column of } A \text{ is zero}\}.$$

Since the first column of $\iota(x)$ is the coordinate vector of $x \in K$, one has $\iota(K) \cap W = \{0\}$, and therefore

$$\mathrm{M}_3(\mathbb{Q}) = \iota(K) \oplus W.$$

Define

$$\Lambda_{K,\alpha} := \mathbb{Z}(\alpha I) \oplus \mathbb{Z}T \oplus \mathbb{Z}T^2 \oplus \bigoplus_{r=1}^{3} \big(\mathbb{Z}E_{r2} \oplus \mathbb{Z}E_{r3}\big).$$

The associated determinant form is

$$F_{\Lambda_{K,\alpha}}(x, y) = \det \begin{pmatrix} \alpha x_1 & 2x_3 + y_{12} & 2x_2 + y_{13} \\ x_2 & \alpha x_1 + y_{22} & 2x_3 + y_{23} \\ x_3 & x_2 + y_{32} & \alpha x_1 + y_{33} \end{pmatrix}.$$

Since

$$\frac{\det(\alpha I)}{\det(T)} = \frac{\alpha^3}{2} \notin \mathbb{Q},$$

the lattice is not determinant-rational, so Theorem 1.1 applies. Moreover, the rank-six submodule

$$\bigoplus_{r=1}^{3}\big(\mathbb{Z}E_{r2}\oplus\mathbb{Z}E_{r3}\big)$$

consists entirely of matrices with zero first column. Since $d=3$, Theorem 3.15 applies, and standard lattice-point counting in this submodule gives

$$c^{\mathrm{sing}}_{\Lambda_{K,\alpha}}>0.$$

Example 13.5 (A cubic reduced-norm form)**.** Let

$$K=\mathbb{Q}(\theta),\qquad \theta^3+\theta^2-2\theta-1=0.$$

This is the real cyclic cubic subfield of $\mathbb{Q}(\zeta_7)$. Let σ generate $\mathrm{Gal}(K/\mathbb{Q})$, and consider the cyclic algebra

$$A=(K/\mathbb{Q},\sigma,2)=K\oplus Ku\oplus Ku^2,$$

with multiplication determined by

$$u^3=2,\qquad ux=\sigma(x)u\quad (x\in K).$$

The ring of integers of K is $\mathbb{Z}[\theta]$, and the reduction of the defining polynomial modulo 2 is X^3+X^2+1, which is irreducible over $\mathbb{F}_2$. Hence 2 is inert in K. At the place above 2, norms from K have valuation divisible by 3, so 2 is not a norm. The norm criterion for cyclic algebras therefore shows that A is not split; since its degree is the prime 3, it is a division algebra. We use here the standard facts on cyclic algebras and reduced norms; see [18].

The Brauer group of $\mathbb{R}$ has exponent 2, whereas the class of A has exponent dividing 3. Thus A splits over $\mathbb{R}$. More concretely, the splitting

$$A\otimes_{\mathbb{Q}}K\simeq \mathrm{M}_3(K)$$

yields, after choosing a real embedding of K, an injective $\mathbb{Q}$-algebra homomorphism

$$\iota: A\hookrightarrow \mathrm{M}_3(K)\hookrightarrow \mathrm{M}_3(\mathbb{R}).$$

In particular, all entries of $\iota(A)$ are algebraic.

Let $\alpha\in\overline{\mathbb{Q}}\cap\mathbb{R}$ satisfy $\alpha^3\notin\mathbb{Q}$, and define

$$\Lambda_{A,\alpha}:=\mathbb{Z}(\alpha I)\oplus\mathbb{Z}\iota(\theta)\oplus\mathbb{Z}\iota(\theta^2)\oplus\bigoplus_{r=0}^{2}\big(\mathbb{Z}\iota(\theta^r u)\oplus\mathbb{Z}\iota(\theta^r u^2)\big).$$

The associated determinant form, with the reduced norm extended to $A\otimes_{\mathbb{Q}}\mathbb{R}$, is

$$\begin{aligned}F_{A,\alpha}(x)=\mathrm{Nrd}_A\big(&\alpha x_1+x_2\theta+x_3\theta^2+x_4u+x_5\theta u\\&+x_6\theta^2u+x_7u^2+x_8\theta u^2+x_9\theta^2u^2\big).\end{aligned}$$

Since $\mathrm{Nrd}_A(u)=2$,

$$\frac{\det(\alpha I)}{\det(\iota(u))}=\frac{\alpha^3}{2}\notin\mathbb{Q}.$$

Thus $\Lambda_{A,\alpha}$ is not determinant-rational, and Theorem 1.1 applies.

This example has no rank-six submodule on which the determinant vanishes identically. Indeed, let

$$W_0 := \mathbb{Q}\iota(\theta) \oplus \mathbb{Q}\iota(\theta^2) \oplus \bigoplus_{r=0}^{2} \big(\mathbb{Q}\iota(\theta^r u) \oplus \mathbb{Q}\iota(\theta^r u^2)\big).$$

Then $\dim_{\mathbb{Q}} W_0 = 8$ and $W_0 \subset \iota(A)$. Since $\alpha \notin \mathbb{Q}$ and the scalar matrices in $\iota(A)$ form its center $\mathbb{Q}I$, one has

$$\mathbb{Q} \cdot (\alpha I) \cap W_0 = \{0\}.$$

Consequently,

$$\mathbb{Q}\Lambda_{A,\alpha} = \mathbb{Q} \cdot (\alpha I) \oplus W_0.$$

Suppose that a rank-six submodule of $\Lambda_{A,\alpha}$ were contained in the singular locus, and let V be its six-dimensional $\mathbb{Q}$-span. The restriction of the determinant to V vanishes on a full lattice, and therefore vanishes identically. Moreover,

$$\dim_{\mathbb{Q}}(V \cap W_0) \geq 6 + 8 - 9 = 5.$$

Choose $0 \neq q \in V \cap W_0$. Since $q \in \iota(A)$,

$$0 = \det(q) = \mathrm{Nrd}_A(q),$$

contradicting the fact that every nonzero element of the division algebra A is invertible. This proves the claim. Since $d = 3$, the isotropic noncoincidence condition is vacuous. By the finite decomposition in Lemma 10.7, the absence of a six-dimensional rational isotropic subspace eliminates every main piece in the singular asymptotic; the remaining pieces contribute only $O(T^3)$. Consequently,

$$c^{\mathrm{sing}}_{\Lambda_{A,\alpha}} = 0.$$

Example 13.6 (Diagonal lattices satisfy isotropic noncoincidence)**.** Let $(\lambda_{ij})_{1 \leq i,j \leq d} \in (\mathbb{R}^\times)^{d^2}$, and set

$$\Delta_{\boldsymbol{\lambda}} := \bigoplus_{i,j=1}^{d} \lambda_{ij}\mathbb{Z}E_{ij} < \mathrm{M}_d(\mathbb{R}).$$

Equivalently, $\Delta_{\boldsymbol{\lambda}} = g\mathbb{Z}^{d^2}$ under the standard coordinate identification $\mathrm{M}_d(\mathbb{R}) \simeq \mathbb{R}^{d^2}$, where g is diagonal. Then $\Delta_{\boldsymbol{\lambda}}$ satisfies the isotropic noncoincidence condition.

Proof. We prove the column statement; the row statement is identical. Write

$$\Lambda_j := \bigoplus_{i=1}^{d} \lambda_{ij}\mathbb{Z}e_i < \mathbb{R}^d$$

for the lattice formed by the j-th column coordinates. If $U < \mathbb{R}^d$, then

$$\mathcal{L}(U) = \{X \in \mathrm{M}_d(\mathbb{R}) : \mathrm{Col}(X) \subset U\}$$

is $\Delta_{\boldsymbol{\lambda}}$-rational if and only if U is Λ_j-rational for every $1 \le j \le d$. Indeed,

$$\Delta_{\boldsymbol{\lambda}} \cap \mathcal{L}(U) = \bigoplus_{j=1}^{d} (\Lambda_j \cap U),$$

under the decomposition of a matrix into its columns.

Now suppose that U is proper and $\dim U \le d-2$. Choose a coordinate vector $e_r \notin U$. Since the line $\mathbb{R}e_r$ is rational with respect to every Λ_j, the subspace

$$U' := U + \mathbb{R}e_r$$

is again Λ_j-rational for every j. Hence $\mathcal{L}(U')$ is a proper $\Delta_{\boldsymbol{\lambda}}$-rational column-isotropic subspace which properly contains $\mathcal{L}(U)$.

Therefore no proper $\Delta_{\boldsymbol{\lambda}}$-rational column-isotropic subspace of dimension kd with $k \le d-2$ is maximal among proper rational column-isotropic subspaces. Thus the column isotropic noncoincidence condition is vacuous. The same argument applied to the row coordinate lattices proves the row condition. Hence $\Delta_{\boldsymbol{\lambda}}$ satisfies the isotropic noncoincidence condition. □

For example, the lattice

$$\Delta_{\sqrt{2}} := \big\{(x_{ij}) \in \mathrm{M}_d(\mathbb{R}) : x_{11} \in \sqrt{2}\,\mathbb{Z},\ x_{ij} \in \mathbb{Z} \text{ for } (i,j) \neq (1,1)\big\}.$$

satisfies the isotropic noncoincidence condition.

Example 13.7 (A non-vacuous $d=4$, $k=2$ noncoincidence example)**.** We give an example showing that the isotropic noncoincidence condition is not vacuous in the first genuinely new case $d=4$, $k=2$.

Let $\mathsf{C} = \mathbb{R}^4$, $\mathsf{R} = (\mathbb{R}^4)^*$ and write $\mathrm{M}_4(\mathbb{R}) = \mathsf{C} \otimes \mathsf{R}$. Let $C_1 := \langle e_1, e_2\rangle$ and $C_2 := \langle e_3, e_4\rangle$. We construct a lattice through its dual. For $i = 1,2$ and $1 \le j \le 4$, choose full lattices $L_{i,j} < C_i$ with algebraic basis vectors, and set

$$\Gamma_i := \bigoplus_{j=1}^{4} L_{i,j} \otimes e_j^* \subset C_i \otimes \mathsf{R}.$$

Define

$$\Delta^* := \Gamma_1 \oplus \Gamma_2 \subset (C_1 \oplus C_2) \otimes \mathsf{R} = \mathrm{M}_4(\mathbb{R}), \qquad \Delta := (\Delta^*)^*.$$

Then Δ is a lattice with algebraic entries.

We choose the lattices $L_{i,j}$ so that, for each $i = 1,2$, there is no line in C_a which is rational with respect to all four lattices $L_{a,1}, \dots, L_{a,4}$. For instance, in $C_1 \simeq \mathbb{R}^2$, one may start with

$$L_{1,1} = \mathbb{Z}e_1 \oplus \mathbb{Z}e_2, \qquad L_{1,2} = \mathbb{Z}e_1 \oplus \mathbb{Z}\sqrt{2}\,e_2,$$

so that the only common rational lines for $L_{1,1}$ and $L_{1,2}$ are the two coordinate axes. Choosing $L_{1,3}$ with algebraic basis vectors so that neither coordinate axis is rational for $L_{1,3}$, and then choosing $L_{1,4}$ generically, gives the desired property. The same construction is made in C_2. Consider the two column-isotropic subspaces $V_1 := C_2 \otimes \mathsf{R}$, $V_2 := C_1 \otimes \mathsf{R}$ of $\mathrm{M}_4(\mathbb{R})$. Their orthogonal complements are

$$V_1^{\perp} = C_1 \otimes \mathsf{R}, \qquad V_2^{\perp} = C_2 \otimes \mathsf{R},$$

and hence

$$V_1^{\perp} \cap \Delta^* = \Gamma_1, \qquad V_2^{\perp} \cap \Delta^* = \Gamma_2.$$

The choice above implies that V_1 and V_2 are maximal proper Δ-rational column-isotropic subspaces of dimension 8. Indeed, if V_1 were contained in a larger proper Δ-rational column-isotropic subspace, then $V_1^{\perp} \cap \Delta^*$ would contain a Δ^*-rational subspace of the form $\ell \otimes \mathsf{R}$, where $\ell < C_1$ is a line. This would mean that ℓ is rational with respect to every $L_{1,j}$, contrary to the construction. The argument for V_2 is the same.

We now arrange that these two maximal pieces are noncoincident. Choose the lattices so that

$$\mathrm{covol}(L_{1,1}) = \mathrm{covol}(L_{2,1}) = 1, \qquad \mathrm{covol}(L_{1,2}) = \sqrt{2}, \qquad \mathrm{covol}(L_{2,2}) = \sqrt{3}.$$

Suppose, for contradiction, that the column noncoincidence condition fails for V_1 and V_2. By the dual formulation of the condition, there exists $g \in \mathrm{GL}_4(\mathbb{R})$ such that

$$g(V_1^{\perp} \cap \Delta^*) \quad \text{is commensurable with} \quad V_2^{\perp} \cap \Delta^*.$$

Since $V_1^{\perp} = C_1 \otimes \mathsf{R}$ and $V_2^{\perp} = C_2 \otimes \mathsf{R}$, the map g sends C_1 to C_2. Let

$$A := g|_{C_1} : C_1 \to C_2.$$

Because g acts by left multiplication, the same linear map A acts on each column. Therefore, for every $1 \leq j \leq 4$,

$$AL_{1,j} \quad \text{is commensurable with} \quad L_{2,j}.$$

Taking covolumes gives

$$|\det A| \, \frac{\mathrm{covol}(L_{1,j})}{\mathrm{covol}(L_{2,j})} \in \mathbb{Q}^{\times} \qquad (1 \leq j \leq 4).$$

For $j = 1$, this implies

$$|\det A| \in \mathbb{Q}^{\times}.$$

For $j = 2$, it implies

$$|\det A| \frac{\sqrt{2}}{\sqrt{3}} \in \mathbb{Q}^{\times},$$

which is impossible. Hence the two 8-dimensional maximal column-isotropic subspaces V_1 and V_2 do not violate the noncoincidence condition.

Finally, by choosing the remaining lattices $L_{a,j}$ generically with algebraic basis vectors, one may ensure that there are no other accidental maximal 8-dimensional column- or row-isotropic rational subspaces, except for those forced by the construction above and their row analogues if one chooses to include them. Since the possible accidental coincidences are described by countably many proper algebraic conditions on the chosen basis vectors, such algebraic choices exist. For these choices, Δ satisfies the isotropic noncoincidence condition, and the $k = 2$ column part of the condition is genuinely non-vacuous.

Appendix A. Analytic stability of negative moments

In this appendix we prove the following lemma which is a main ingredient of the proof of Lemma 4.5.

Lemma A.1. *Let $B \subset \mathbb{R}^n$ be a bounded semialgebraic open set, let $J = [0, \delta)$, and let*

$$H : J \times B \to [0, \infty)$$

be the restriction of a real-analytic globally subanalytic function defined on a neighborhood of $J \times \overline{B}$. Suppose that, for some $\tau > 0$,

$$H(0, \cdot)^{-\tau} \in L^1(B).$$

Then, for every $0 < \sigma < \tau$, there exists $\delta' > 0$ such that

$$\sup_{0 \le t < \delta'} \int_B H(t, x)^{-\sigma}\, dx < \infty.$$

Proof. The proof uses standard consequences of boundary-compatible rectilinearization and the monomial integrability criterion of Cluckers–Miller. Apply parameterized rectilinearization to H, the parameter function t, and a finite collection of semialgebraic functions defining the boundary of B; in doing so, apply the preparation recursively on the frontier $\{t = 0\}$. This is the boundary-compatible form of [9, Theorem 1.5]. After a finite partition and power substitutions, each full-dimensional chart whose closure meets $\{t = 0\}$ has coordinates $y = (y_1, \ldots, y_r) \in (0, 1)^r$ in which

$$t \circ \pi = y^\beta, \qquad H \circ \pi = y^\alpha u(y), \qquad |\det D\pi(y)| = y^\gamma v(y). \tag{A.1}$$

Here $\alpha, \beta \in \mathbb{N}^r$, $\gamma \in \mathbb{Z}^r$, $\beta \neq 0$, and u, v are positive analytic units bounded above and below away from zero. A positive unit in the expression for $t \circ \pi$ has been absorbed into one of the coordinates with $\beta_i > 0$. Extra coordinates ranging in a fixed compact subset of $(0, 1]$ are harmless and are suppressed from the notation. Lower-dimensional charts that dominate the parameter interval have zero measure in every fiber, while charts contained in $\{t = 0\}$ contribute only to the central fiber.

We need one standard consequence of carrying out the rectilinearization compatibly with $\{t = 0\}$. For each coordinate divisor $E_i = \{y_i = 0\}$ other than the strict transform of $\{t = 0\}$, there is a corresponding divisor over the central fiber for which, writing $\kappa_i^{(0)}$ for the central-fiber Jacobian exponent and $\alpha_i^{(0)}$ for the order of $H(0, \cdot)$,

$$\kappa_i^{(0)} + 1 = \gamma_i + 1 - \beta_i, \qquad \alpha_i^{(0)} \ge \alpha_i. \tag{A.2}$$

The first identity is the usual adjunction calculation for the fiber $t = 0$; the second inequality holds because restriction to $t = 0$ can only increase the order of vanishing. These relations also follow directly from the recursive construction: when a coordinate center is blown up, the new Jacobian and parameter orders add, and the same blow-up on the central fiber lowers the Jacobian order by precisely the order of t. Thus the quantity $(\gamma_i + 1) - \beta_i$ is exactly the corresponding central-fiber log-Jacobian exponent.

Applying the monomial integrability criterion [9, Corollary 5.5] on the resolved central fiber gives

$$\kappa_i^{(0)} - \tau\alpha_i^{(0)} > -1.$$

Together with (A.2), this implies

$$d_i(\tau) := \gamma_i + 1 - \beta_i - \tau\alpha_i > 0 \tag{A.3}$$

for every coordinate divisor other than the strict transform of $\{t = 0\}$. For the strict-transform coordinate, one has

$$(\beta_i, \gamma_i, \alpha_i) = (1, 0, 0), \tag{A.4}$$

so $d_i(\tau) = 0$. Indeed, $H(0, \cdot)$ is not identically zero on any nonempty open subset of B, and therefore H has order zero along the generic point of the strict transform of $\{t = 0\}$.

Since $\alpha_i \geq 0$ and $\sigma < \tau$,

$$d_i(\sigma) = d_i(\tau) + (\tau - \sigma)\alpha_i > 0 \tag{A.5}$$

for every coordinate other than the possible strict-transform coordinate; for that coordinate, $d_i(\sigma) = 0$. Thus among the coordinates with $\beta_i > 0$, at most one has $d_i(\sigma) = 0$.

It remains to estimate the fiber integral in a chart. Put

$$\mathscr{Q} := \{i : \beta_i = 0\}, \qquad \mathcal{V} := \{i : \beta_i > 0\}.$$

For $i \in \mathscr{Q}$, (A.5) reads $\gamma_i - \sigma\alpha_i > -1$. Hence integration in the horizontal variables is uniformly bounded by Corollary 5.5 of [9]. After performing those integrations, the coarea formula bounds the remaining contribution by a constant multiple of

$$\int_{\{y\in(0,1)^{\mathcal{V}}: y^\beta = t\}} y^{\gamma - \sigma\alpha} \frac{d\mathscr{Q}^{|\mathcal{V}|-1}(y)}{|\nabla(y^\beta)|}, \tag{A.6}$$

where only the coordinates indexed by $\mathcal{V}$ occur in the monomials.

Set $y_i = e^{-w_i}$ and $L = -\log t$. The coarea formula, or equivalently the identity

$$\delta\big(t - e^{-\beta\cdot w}\big) = t^{-1}\delta\big(L - \beta \cdot w\big),$$

shows that (A.6) is bounded by a constant multiple of

$$\int_{\substack{w_i \geq 0 \ (i\in\mathcal{V}) \\ \sum_{i\in\mathcal{V}} \beta_i w_i = L}} \exp\left(-\sum_{i\in\mathcal{V}} d_i(\sigma) w_i\right) d\lambda_L(w), \tag{A.7}$$

where $d\lambda_L$ is Euclidean measure on the indicated hyperplane; the normalizing factor depends only on β.

If all $d_i(\sigma)$ are positive, let

$$c := \min_{i \in \mathcal{V}} \frac{d_i(\sigma)}{\beta_i} > 0.$$

On the hyperplane in (A.7),

$$\sum_{i \in \mathcal{V}} d_i(\sigma) w_i \geq cL.$$

The volume of the corresponding simplex is $O((1+L)^{|\mathcal{V}|-1})$, so the integral is

$$O\big((1+L)^{|\mathcal{V}|-1} e^{-cL}\big),$$

uniformly for $L \geq 0$. If a strict-transform coordinate i_0 is present, solve the hyperplane equation for w_{i_0}. Since $d_{i_0}(\sigma) = 0$, enlarging the remaining domain to the full positive orthant gives

$$(\text{A.7}) \ll \prod_{\substack{i \in \mathcal{V} \\ i \neq i_0}} \int_0^\infty e^{-d_i(\sigma) w_i} \, dw_i < \infty,$$

again uniformly in L. Summing over the finitely many charts proves the lemma. □

References

[1] Kaan Akin, David A. Buchsbaum, and Jerzy Weyman. Schur functors and Schur complexes. *Adv. Math.*, 44(3):207–278, 1982.

[2] Yves Benoist and Jean-Francois Quint. Random walks on finite volume homogeneous spaces. *Inv. Math.*, 187(1):37–59, 2012.

[3] B. J. Birch. Forms in many variables. *Proc. Roy. Soc. London Ser. A*, 265:245–263, 1962.

[4] Jacek Bochnak, Michel Coste, and Marie-Francoise Roy. *Real Algebraic Geometry*, volume 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3)*. Springer, Berlin, Heidelberg, 1998.

[5] A. Borel and Harish-Chandra. Arithmetic subgroups of algebraic groups. *Ann. of Math. (2)*, 75:485–535, 1962.

[6] Yu. A. Brudnyi and M. I. Ganzburg. On an extremal problem for polynomials in n variables. *Izv. Akad. Nauk SSSR Ser. Mat.*, 37:344–355, 1973.

[7] J.W.S. Cassels. *An Introduction to the Geometry of Numbers*. Springer-Verlag, Berlin-New York, 1971. Second printing, corrected, Die Grundlehren der mathematischen Wissenschaften, Band 99.

[8] Raf Cluckers and Daniel J. Miller. Stability under integration of sums of products of real globally subanalytic functions and their logarithms. *Duke Math. J.*, 156(2):311–348, 2011.

[9] Raf Cluckers and Daniel J. Miller. Lebesgue classes and preparation of real constructible functions. *J. Funct. Anal.*, 264(7):1599–1642, 2013.

[10] S. G. Dani and G. A. Margulis. Limit distributions of orbits of unipotent flows and values of quadratic forms. In *I. M. Gel'fand Seminar*, volume 16 of *Advances in Soviet Mathematics*, pages 91–137. American Mathematical Society, Providence, RI, 1993.

[11] H. Davenport. On a principle of lipschitz. *J. London Math. Soc.*, 26:179–183, 1951.

[12] E. B. Dynkin. Maximal subgroups of the classical groups. *Trudy Moskov. Mat. Obsc*, 1:39–166, 1952.

[13] A. Eskin, G. Margulis, and S. Mozes. Upper bounds and asymptotics in a quantitative version of the oppenheim conjecture. *Ann. of Math.*, 147(1):93–141, 1998.

[14] A. Eskin, G. Margulis, and S. Mozes. Quadratic forms of signature $(2, 2)$ and eigenvalue spacings on rectangular 2-tori. *Ann. of Math. (2)*, 161(2):679–725, 2005.

[15] William Fulton. *Young tableaux*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1997. With applications to representation theory and geometry.

[16] William Fulton. *Young Tableaux: With Applications to Representation Theory and Geometry*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1997.

[17] William Fulton and Joe Harris. *Representation Theory: A First Course*, volume 129 of *Graduate Texts in Mathematics*. Springer, New York, 1991.

[18] Philippe Gille and Tamas Szamuely. *Central Simple Algebras and Galois Cohomology*, volume 101 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2006.

[19] James A. Green. *Polynomial representations of* GL_n, volume 830 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1980.

[20] Y. R. Katznelson. Integral matrices of fixed rank. *Proc. Amer. Math. Soc.*, 120(3):667–675, 1994.

[21] Wooyeon Kim. Moments of margulis functions and indefinite ternary quadratic forms. *arXiv:2403.16563*, 2024.

[22] Wooyeon Kim and Hee Oh. Quantitative Oppenheim in signature $(2, 2)$ via determinant values. *Preprint*, 2026.

[23] D. Y. Kleinbock and G. A. Margulis. Flows on homogeneous spaces and Diophantine approximation on manifolds. *Ann. of Math. (2)*, 148(1):339–360, 1998.

[24] Dmitry Kleinbock. An extension of quantitative nondivergence and applications to diophantine exponents. *Transactions of the American Mathematical Society*, 360:6497–6523, 2008.

[25] Marvin Marcus and B. N. Moyls. Linear transformations on algebras of matrices. *Canadian Journal of Mathematics*, 11:61–66, 1959.

[26] G. A. Margulis. Quadratiques indéfinies et flots unipotents sur les espaces homogènes. *C. R. Acad. Sci. Paris Sér. I Math.*, 304(10):249–253, 1987.

[27] G. A. Margulis. Discrete subgroups and ergodic theory. In *Number Theory, Trace Formulas and Discrete Groups (Oslo, 1987)*, pages 377–398. Academic Press, Boston, MA, 1989.

[28] M. Ratner. Raghunathan's topological conjecture and distributions of unipotent flows. *Duke Math. J.*, 63(1):235–280, 1991.

[29] Marina Ratner. On Raghunathan's measure conjecture. *Ann. of Math. (2)*, 134(3):545–607, 1991.

[30] Nimish A. Shah. Limit distributions of expanding translates of certain orbits on homogeneous spaces. *Proc. Indian Acad. Sci. Math. Sci.*, 106(2):105–125, 1996.

[31] William C. Waterhouse. Twisted forms of the determinant. *Journal of Algebra*, 86(1):60–75, 1984.

[32] Robert J. Zimmer. *Ergodic Theory and Semisimple Groups*, volume 81 of *Monographs in Mathematics*. Birkhäuser, Boston, 1984.

Korea Institute for Advanced Study, Seoul, Republic of Korea
Email address: `wooyeonkim@kias.re.kr`

Department of Mathematics, Yale University, New Haven, CT
Email address: `hee.oh@yale.edu`