GEOMETRY OF NUMBERS WITH APPLICATIONS TO NUMBER THEORY

PETE L. CLARK

Contents

Thanks to Lauren Huckaba for a careful reading of §6, including spotting some typos and fleshing out some details in the proof of the Minkowski- Hlawka Theorem.

1. Lattices in Euclidean Space

Fix a positive integer N, and consider N-dimensional Euclidean space $\mathbb{R}^N$. In particular, under addition $\mathbb{R}^N$ has the structure of a locally compact topological group. A **vector group** is a topological group G isomorphic to a subgroup of $(\mathbb{R}^N, +)$ for some N. In particular, $\mathbb{R}^N$ is a vector group.

1.1. **Discrete vector groups.**

At the other extreme are the **discrete subgroups**: those for which the induced subspace topology is the discrete topology.

Exercise 1.1: Show that for a subgroup $G \subset \mathbb{R}^N$ the following are equivalent:
(i) The infimum of the lengths of all nonzero elements of G is positive.
(ii) G is discrete.

Proposition 1.1. *Let G be a Hausdorff topological group and H a locally compact subgroup. Then H is closed in G. In particular, every discrete subgroup of a Hausdorff group is closed.*

Proof. Let K be a compact neighborhood of the identity in H. Let U be an open neighborhhod of the identity in G such that $U \cap H \subset K$. Let $x \in \overline{H}$. Then there is a neighborhood V of x such that $V^{-1}V \subset U$, so then

$$(V \cap H)^{-1}(V \cap H) \subset K.$$

Since $x \in \overline{H}$, there exists $y \in V \cap H$, and then $V \cap H \subset yK$. Since for every neighborhood W of x, $W \cap V$ is also a neighborhood of x and thus $W \cap V \cap H \neq \emptyset$, $x \in \overline{V \cap H}$. Since yK is compact in the Hausdorff space H, it is closed and thus $x \in \overline{V \cap H} \subset \overline{yK} = yK \subset H$. So H is closed. □

Example 1.2: For every $0 \leq n \leq N$, there is a discrete subgroup of $\mathbb{R}^N$ isomorphic as a group to $\mathbb{Z}^n$. This is almost obvious: the subgroup $\mathbb{Z}^N$ of $\mathbb{R}^N$ – what we call often call the **standard integer lattice** – is discrete, and hence so too are the subgroups $\mathbb{Z}^n = \mathbb{Z}^N \cap (\mathbb{R}^n \times 0^{N-n})$. Thus for all $n \leq N$, there are discrete subgroups of $\mathbb{R}^N$ which are, as abstract groups, free abelian of rank n.

In fact the converse is also true: every discrete subgroup of $\mathbb{R}^N$ is free abelian of rank at most N. But this is not obvious! The following exercise drives this home.

Exercise 1.3: a) Show that, as abstract groups, $(\mathbb{R}^N, +) \cong (\mathbb{R}, +)$.[1]
b) Deduce that for all $n, N \in \mathbb{Z}^+$, $\mathbb{R}^N$ admits a subgroup $G \cong \mathbb{Z}^n$.
c) Show in fact that $\mathbb{R}^N$ admits a subgroup which is free abelian of rank κ for every cardinal number $\kappa \leq \#\mathbb{R}$.

Define the **real rank** $\mathfrak{r}(G)$ of a vector group $G \subset \mathbb{R}^N$ to be the maximal cardinality of an $\mathbb{R}$-linearly independent subset of G, so $0 \leq \mathfrak{r}(G) \leq N$. For instance, the discrete subgroup $\mathbb{Z}^n \subset \mathbb{Z}^N \subset \mathbb{R}^N$ of Example 1.1 above has real rank n.

Theorem 1.2. *Let G be a discrete subgroup of $\mathbb{R}^N$, of real rank r. There are $\mathbb{R}$-linearly independent vectors $v_1, \ldots, v_r$ such that $G = \langle v_1, \ldots, v_r \rangle_{\mathbb{Z}}$.*

Proof. By Proposition 1.1, G is closed.First observe that $r = 0 \iff G = \{0\}$ and this is a trivial case. Henceforth we assume $0 < r \leq N$. By definition of real rank, there are $e_1, \ldots, e_r \in G$ which are $\mathbb{R}$-linearly independent. Let $P = \{\sum_{i=1}^r x_i e_i \mid e_i \in [0,1]\}$ be the corresponding paralleletope. Thus $G \cap P$ is a closed, discrete subspace of a compact set, hence finite. Let $x \in G$. Since r is the real rank of G, there are $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ such that $x = \sum_{i=1}^r \lambda_i e_i$. For $j \in \mathbb{Z}$, put

$$x_j = jx - \sum_{i=1}^r \lfloor j\lambda_i \rfloor e_i.$$

Thus

$$x_j = \sum_{i=1}^r (j\lambda_i - \lfloor j\lambda_i \rfloor) e_i,$$

so $x_j \in P \cap G$. Since $x = x_1 + \sum_{i=1}^r \lfloor \lambda_i \rfloor e_i$, we seee that G is generated as a $\mathbb{Z}$-module by $G \cap P$ hence is finitely generated. Further, since $G \cap P$ is finite and $\mathbb{Z}$ is infinite, there are distinct $j, k \in \mathbb{Z}$ such that $x_j = x_k$. Then

$$(j-k)\lambda_i = \lfloor j\lambda_i \rfloor - \lfloor k\lambda_i \rfloor,$$

[1]We maintain our convention that N is an arbitrary, "fixed" positive integer.

so $\lambda_i \in \mathbb{Q}$. Thus G is generated as a $\mathbb{Z}$-module by a finite number of $\mathbb{Q}$-linear combinations of the e_i's. Let d be a common denominator of these coefficients, so $dG \subset \sum_{i=1}^r \mathbb{Z}e_i$. By the sturcture theory of finitely generated modules over a PID, there is a $\mathbb{Z}$-basis $f_1, \ldots, f_r$ of $\sum_{i=1}^r \mathbb{Z}e_i$ and integers $\alpha_1, \ldots, \alpha_r$ such that $dG = \langle \alpha_1 f_1, \ldots, \alpha_r f_r \rangle$. The free rank of dG is the same as the free rank of G, which is at least r since $e_1, \ldots, e_r \in G$. It follows that the free rank of G is exactly r and the integers α_i are all nonzero. It follows that dG is generated as a $\mathbb{Z}$-module by the $\mathbb{R}$-linearly indepenent vectors $f_1, \ldots, f_r$, hence G is generated as a $\mathbb{Z}$-module by the $\mathbb{R}$-linearly independent vectors $v_1 = \frac{f_1}{d}, \ldots, v_r = \frac{f_r}{d}$. □

Exercise 1.4: Deduce from Theorem 1.2 the following purely algebraic result: any subgroup of a free abelian group of finite rank n is free abelian of rank at most n.

We define a **lattice** in $\mathbb{R}^N$ to be a discrete subgroup which is, as an abstract group, free abelian of rank N. (Note that what we call a "lattice" is sometimes called a "full lattice" or a "lattice of full rank" by other authors.) By Theorem 1.2, a subgroup of $\mathbb{R}^N$ is a lattice iff it is the $\mathbb{Z}$-span of an $\mathbb{R}$-basis $v_1, \ldots, v_N$ for $\mathbb{R}^N$: we refer to $\{v_1, \ldots, v_N\}$ simply as a **basis** for Λ.

Exercise 1.5: Let $\Lambda \subset \mathbb{R}^N$ be a lattice, and let $S \subset \Lambda$ be a subset. Show that S is $\mathbb{Z}$-linearly independent iff it is $\mathbb{R}$-linearly independent.

1.2. **Hermite and Smith Normal Forms.**

In general, the notions of Hermite and Smith normal forms belong to the structure theory of finitely generated modules over a PID. We restrict ourselves to the case of direct relevance here: let $N \in \mathbb{Z}^+$, and let $M \in M_N(\mathbb{Z})$.

The matrix $M = (m_{ij})$ is in **Hermite normal form (HNF)** if:
(HNF1) M is upper triangular: $m_{i,j} = 0$ for all $i > j$,
(HNF2) $m_{i,i} > 0$ for all $1 \leq i \leq N$, and
(HNF3) For all $1 \leq i < j \leq N$, $0 \leq m_{ij} < m_{ii}$.

The matrix $M = (m_{ij})$ is in **Smith normal form (SNF)** if
(SNF1) M is diagonal: $m_{ij} = 0$ for all $i \neq j$,
(SNF2) $m_{i,i} \geq 0$ for all $1 \leq i \leq N$, and
(SNF3) Forall $1 \leq i < N$, $m_{ii} \mid m_{i+1,i+1}$.

Theorem 1.3. *(Hermite) Let $A \in M_N(\mathbb{Z})$ with $\det A \neq 0$. Then there is a unique matrix M in Hermite normal form such that $M = AU$ for some $U \in \mathrm{GL}_N(\mathbb{Z})$.*

Proof. See [Coh, §2.4.2] for a constructive proof, i.e., an algorithm for putting A in Hermite Normal Form together with a proof of its correctness. □

Exercise: Show that Theorem 1.3 is equivalent to the following statement about $\mathbb{Z}$-modules: let $\Lambda \subset \mathbb{Z}^N$ be a free abelian group of rank N. Then there is a $\mathbb{Z}$-basis $v_1, \ldots, v_N$ of Λ and $M = (m_{ij}) \in M_N(\mathbb{Z})$ in HNF such that

$$v_1 = m_{11}e_1 + m_{12}e_2 + \ldots + m_{1N}e_N,$$
$$v_2 = m_{22}e_2 + m_{23}e_3 + \ldots + m_{2N}e_N,$$
$$\vdots$$

$$v_N = m_{NN} e_N.$$

Theorem 1.4. *(Smith) Let $A \in M_N(\mathbb{Z})$. Then there is a unique matrix M in Smith normal form such that $M = VAU$ for some $U, V \in \mathrm{GL}_N(\mathbb{Z})$.*

Proof. In [J1] a more general result is given (for not necessarily square matrices with coefficients in an arbitrary PID). [Coh, §2.4.4] gives a constructive proof when $\det A \neq 0$, which can be easily adapted to the singular case. □

Exercise: Show that Theorem 1.4 is equivalent to the following statement about $\mathbb{Z}$-modules: let $\Lambda_1 \cong \mathbb{Z}^N$, and let Λ_2 be a subgroup. Then there is a $\mathbb{Z}$-basis $v_1, \ldots, v_N$ for Λ_1 and positive integers $d_1 \mid d_2 \mid \ldots \mid d_N$ such that $d_1 v_1, \ldots, d_N v_N$ generates Λ_2, and restricting to the $d_i v_i$'s with $d_i \neq 0$ gives a $\mathbb{Z}$-basis for Λ_2.

Corollary 1.5. *Let p be a prime number, $n, N \in \mathbb{Z}^+$, and $V = \bigoplus_{i=1}^N \mathbb{Z}/p^n\mathbb{Z}$. Let H be a subgroup of V.*
a) There are unique natural numbers $0 \leq n_1 \leq n_2 \leq \ldots \leq n_N \leq n$ such that

$$H \cong \bigoplus_{i=1}^N \mathbb{Z}/p^{n_i}\mathbb{Z}.$$

b) We have $V/H \cong \bigoplus_{i=1}^N \mathbb{Z}/p^{n-n_i}\mathbb{Z}$.

Proof. Let Λ_0 be a free abelian group of rank N, and let $\Lambda_2 = p^n\Lambda$. Then $\Lambda_0/\Lambda_2 \cong V$, so by the correspondence principle there is a subgroup $\Lambda_1 \subset \Lambda_0$ such that $\Lambda_1/\Lambda_2 = H$. Apply Smith Normal Form to Λ_1 and Λ_2: there exists a basis $v_1, \ldots, v_N$ of Λ_1 and positive integers $d_1, \ldots, d_N$ with $d_i \mid d_{i+1}$ such that $d_1 v_1, \ldots, d_N v_N$ is a basis of Λ_2, and thus

$$H = \Lambda_1/\Lambda_2 \cong \bigoplus_{i=1}^N \mathbb{Z}/d_i\mathbb{Z}.$$

Since H is a p-group, we may write $d_i = p^{n_i}$, establishing part a).
Next note that since $\Lambda_2 = p^n\Lambda_1$, $\frac{p^{n_1}}{p^n}e_1, \ldots, \frac{p^{n_N}}{p^n}e_N$ is a $\mathbb{Z}$-basis for Λ_0, and thus

$$V/H = (\Lambda_0/\Lambda_2)/(\Lambda_1/\Lambda_2) \cong \Lambda_0/\Lambda_1 \cong \bigoplus_{i=1}^N \mathbb{Z}/p^{n-n_i}\mathbb{Z}.$$

□

1.3. **Fundamental regions, covolumes and sublattices.**

For a group G acting on a space X, a **fundamental region** is a subset $R \subset X$ containing exactly one element from every G-orbit on X. In other words, a fundamental region is precisely the image of a section of the orbit map $X \to G\backslash X$. Thus the translates of any fundamental region partition the space:

$$X = \coprod_{g \in G} gR.$$

In general there are many fundamental regions, and one looks for fundamental regions with nice topological properties. In general, a fundamental region X need be neither open nor closed, so often it is convenient to deal with **closed fundamental regions** $\overline{X}$, even though these are no longer fundamental regions in the strict sense. Thus we say a family of subsets $\{Y_i\}$ of a topological space X is a **tiling** of X if

(T1) $\bigcup_i \overline{Y_i} = X$, and
(T2) For all $i \neq i'$, $Y_i^\circ \cap Y_j^\circ = \varnothing$.

Note also that there is a natural (quotient) map $q : \overline{X} \to G\backslash X$ which is injective on X°. One important consequence of this is that if $\overline{X}$ is compact, so is $G\backslash X$.

Now we come back to earth: let Λ be a lattice in $\mathbb{R}^N$, and view Λ as acting on $\mathbb{R}^N$ via translations. A particularly nice fundamental region can be obtained using any basis $\mathfrak{v} = \{v_1, \ldots, v_N\}$: namely we define the **fundamental parallelopiped**

$$\mathcal{P}(\mathfrak{v}) = \{\alpha_1 v_1 + \ldots + \alpha_N v_N \mid \alpha_i \in [0,1)\}.$$

Exercise: Show that $\mathcal{P}(\mathfrak{v})$ is a fundamental region for the action of Λ on $\mathbb{R}^N$.

As above, it is also natural to consider the closed parallelopipeds

$$\overline{\mathcal{P}}(\mathfrak{v}) = \{\alpha_1 v_1 + \ldots + \alpha_N v_N \mid \alpha_i \in [0,1]\}.$$

Since $\overline{\mathcal{P}}(\mathfrak{v})$ is compact, the quotient space $\mathbb{R}^N/\Lambda$ is compact. In fact we can say much more: the identifications on the boundary of the closed parallelopiped are precisely that of identifying the $\alpha_i = 0$ face with the $\alpha_i = 1$ face for $1 \leq i \leq N$, and thus the quotient space is an N-dimensional torus, i.e., isomorphic as a topological group to the product of N circles.

One can show that under suitable hypotheses, a measure on the space X descends to give a measure on the quotient X/G. One way to do this is to define measures in terms of fundamental regions. For instance, we wish to define the measure of X/G to be the measure of a fundamental region, and for this to make sense we must check (i) that we may always choose a measurable fundamental region and (ii) any two measurable fundamental regions have the same measure. These arguments are carried through in some generality in XXXX.

In the situation of a lattice acting on Euclidean space things are easier: we can restrict as above to fundamental *parallelotopes*. Since for any basis $\mathfrak{v}$ of $\mathbb{R}^N$, the parallelotopes $\mathcal{P}(\mathfrak{v})$ and $\overline{\mathcal{P}(\mathfrak{v})}$ are the images of the paralleltopes (cubes!) associated to the standard orthogonal basis $e_1, \ldots, e_N$ under the matrix $M_\mathfrak{v}$ with columns $v_1, \ldots, v_N$, we have

$$\operatorname{Vol} \mathcal{P}(\mathfrak{v}) = \operatorname{Vol} \overline{\mathcal{P}(\mathfrak{v})} = |\det M_\mathfrak{v}|.$$

Now suppose $\mathfrak{v}$ and $\mathfrak{w}$ are two bases for the same lattice Λ. Then there is $A \in \mathrm{GL}_N(\mathbb{Z})$ with $AM_\mathfrak{v} = M_\mathfrak{w}$, so

$$|\det M_\mathfrak{w}| = |\det A||\det M_\mathfrak{v}| = |\pm 1||\det M_\mathfrak{v}| = |\det M_\mathfrak{v}|.$$

Thus the volume of a fundamental parallelopiped for Λ is independent of the chosen basis for Λ. We call this invariant the **covolume** of Λ and denote it by $\operatorname{Covol} \Lambda$.

1.3.1. *Sublattices and indices.*

If Λ' and Λ are lattices in $\mathbb{R}^N$ we say that Λ' is a **sublattice** of Λ.

Proposition 1.6. *Let Λ be a lattice in $\mathbb{R}^N$, and let $G \subset \Lambda$ be any subgroup. The following are equivalent:*
(i) G is a lattice.
(ii) The index $[\Lambda : G]$ is finite.
(iii) $\mathbb{R}^N/G$ is compact.

Proof. By Exercise 1.4 (or by Smith Normal Form), G is a free abelian group of rank $n \leq N$. By Smith Normal Form, there are bases $v_1, \ldots, v_N$ for Λ, $w_1, \ldots, w_n$ for G and $d_1, \ldots, d_n \in \mathbb{Z}^+$ such that $w_i = d_i v_i$ for $1 \leq i \leq n$. Thus

$$\Lambda/G \cong \mathbb{Z}^{N-n} \oplus \bigoplus_{i=1}^{n} \mathbb{Z}/d_i\mathbb{Z}$$

and

$$\mathbb{R}^N/G \cong \mathbb{R}^{N-n} \oplus (S^1)^n.$$

From these isomorphisms the equivalence of (i), (ii) and (iii) follows immediately. □

Proposition 1.7. *(Index-Covolume Relation) Let Λ' be a sublattice of Λ. Then*

$$\operatorname{Covol} \Lambda' = [\Lambda : \Lambda'] \operatorname{Covol} \Lambda.$$

Proof. A Smith Normal Form argument works; we leave the details to the reader. □

In fact the result of Proposition 1.7 holds much more generally, and here is an argument which works in this generality: let R be a fundamental region for the action of Γ on X, and let Γ' a finite index subgroup of Γ. Let $g_1, \ldots, g_I$ be a set of coset representatives for Γ' in Γ. Then $\bigcup_{i=1}^{I} g_i R$ is a fundamental region for Γ' and

$$\operatorname{Vol} \bigcup_{i=1}^{I} g_i R = \sum_{i=1}^{I} \operatorname{Vol} g_i R = \sum_{i=1}^{I} \operatorname{Vol} R = n \operatorname{Vol} R.$$

1.3.2. *The number of index n sublattices of $\mathbb{Z}^N$.*

For $n, N \in \mathbb{Z}^+$, let $L_N(n)$ denote the number of index n sublattices of $\mathbb{Z}^N$. It is a nice application of the previous material to evaluate $L_N(n)$ in various cases.

First let us establish that $L_N(n)$ is finite in all cases and give an explicit upper bound. The key idea is that if Λ is an index n sublattice of $\mathbb{Z}^N$, then $\mathbb{Z}^N/\Lambda$ is an n-torsion abelian group, hence $\Lambda \supset (n\mathbb{Z})^N$. Put $\overline{\Lambda} = \Lambda/(n\mathbb{Z})^N$, so $\overline{\Lambda} \subset (\mathbb{Z}/n\mathbb{Z})^N$. Thus the index n sublattices of $\mathbb{Z}^N$ correspond bijectively to index n subgroups of $(\mathbb{Z}/n\mathbb{Z})^N$, i.e., to order n^{N-1} subgroups of $(\mathbb{Z}/n\mathbb{Z})^N$. A crude upper bound is the number of n^{N-1}-element subsets of $(\mathbb{Z}/n\mathbb{Z})^N$, so

$$L_N(n) \leq \binom{n^N}{n^{N-1}} \leq 2^{n^N}.$$

Exercise: Show that $L_1(n) = 1$ for all $n \in \mathbb{Z}^+$.

Proposition 1.8. *For any prime number p, $L_N(p) = \frac{p^N - 1}{p - 1}$.*

Proof. By the above analysis, when $n = p$ is prime, we want to count the codimension one $\mathbb{F}_p$-subspaces of $V \cong \mathbb{F}_p^N$. By duality, the codimension one subspaces of V correspond to the one-dimensional subspaces of $V^\vee \cong \mathbb{F}_p^N$. A one-dimensional subspace of $\mathbb{F}_p^N$ is given as the span of any nonzero vector v and $\langle v_1 \rangle = \langle v_2 \rangle$ iff $v_2 = \alpha v_1$ for some $\alpha \in \mathbb{F}_p^\times$. This leads to the count $L_N(p) = \frac{\#\mathbb{F}_p^N - 1}{p-1} = \frac{p^N-1}{p-1}$. □

Proposition 1.9. *The function $L_N : \mathbb{Z}^+ \to \mathbb{Z}^+$ is multiplicative: that is, for $N, n_1, n_2 \in \mathbb{Z}^+$ with $\gcd(n_1, n_2) = 1$, we have*

$$L_N(n_1 n_2) = L_N(n_1) L_N(n_2).$$

Exercise: Prove Proposition 1.9. (Hint: use the Chinese Remainder Theorem.)

By Proposition 1.9, it is enough to evaluate $L_N(p^n)$ for any prime power p^n. Clearly $L_N(p^0) = 1$, and we have already evaluated $L_N(p)$. Let us examine the next case.

Example 1.10. *We will compute $L_2(p^2)$; equivalently by our preliminary analysis, we wish to count subgroups $\overline{\Lambda}$ of $V = (\mathbb{Z}/p^2\mathbb{Z})^2$ of index p^2, and hence also of order p^2. There are two possible group structures for $\overline{\Lambda}$: $\mathbb{Z}/p^2\mathbb{Z}$ and $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$.*

Suppose first that $\overline{\Lambda} \cong \mathbb{Z}/p^2\mathbb{Z}$. Thus $\overline{\Lambda}$ is generated by a single element of maximal order p^2. The number of elements of order p^2 is $\#V - \#V[p] = p^4 - p^2 = p^2(p^2-1)$. To count subgroups rather than generators of subgroups we divide by the number of generators of a subgroup of order p^2, i.e., $\varphi(p^2) = p^2 - p$, and thus the number of cyclic subgroups $\overline{\Lambda}$ is $\frac{p^2(p^2-1)}{p(p-1)} = p(p+1) = p^2 + p$.

Now suppose $\overline{\Lambda} \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$. Then $\overline{\Lambda} \subset V[p]$ and both have order p^2, so $\overline{\Lambda} = V[p]$, i.e., there is exactly one such subgroup. Thus

$$L_2(p^2) = p^2 + p + 1.$$

Lemma 1.11. *Let p be a prime number and $n, N \in \mathbb{Z}^+$. The number of* cyclic *order p^n subgroups of $V = (\mathbb{Z}/p^n\mathbb{Z})^N$ is*

$$G_{1,N}(p^n) = \frac{(p^n)^N - (p^{n-1})^N}{\varphi(p^n)} = \frac{p^{(n-1)N}(p^N - 1)}{p^{n-1}(p-1)} = p^{(N-1)(n-1)} \left(1 + p + \ldots + p^{N-1}\right).$$

Exercise: Prove Lemma 1.11. (Suggestion: adapt Example 1.10 above.)

Theorem 1.12. *For any $n \in \mathbb{Z}^+$, $L_2(n) = \sigma(n) = \sum_{d|n} d$.*

Proof. Step 1: We have already seen that the function L_2 is multiplicative. So too is the divisor sum function, so it suffices to show the result in the case n is a prime power, in which case an equivalent form of the statement to be proved is

$$L_2(p^n) = p^n + p^{n-1} + \ldots + p + 1.$$

Putting $V = (\mathbb{Z}/p^n\mathbb{Z})^2$, we want to count subgroups $\overline{\Lambda}$ of index p^n, equivalently of order p^n. By Corollary 1.5 such a subgroup is of the form $\mathbb{Z}/p^{n-a} \oplus \mathbb{Z}/p^a$ for $0 \le a \le \lceil \frac{n-1}{2} \rceil$: we call these subgroups of **type a**. We claim that the number of subgroups of type a is $p^{n-2a} + p^{n-2a-1}$ unless n is even and $a = \frac{n}{2} = \lceil \frac{n-1}{2} \rceil$, in which case there is exactly one subgroup of type a. In the final case we have $\overline{\Lambda} \subset V[p^{\frac{n}{2}}]$ and both have p^n elements, so we must have $\overline{\Lambda} = V[p^{\frac{n}{2}}]$.

Suppose $a = 0$, so $\overline{\Lambda} \cong \mathbb{Z}/p^n$. Then by Lemma 1.11 the number of type 0 subgroups is $C(2, p^n) = p^{n-1}(p+1) = p^n + p^{n-1}$.

Now suppose $0 < a < \frac{n}{2}$, so $\overline{\Lambda} \cong \mathbb{Z}/p^{n-a}\mathbb{Z} \times \mathbb{Z}/p^a\mathbb{Z}$ and thus $\overline{\Lambda} \subset V[p^{n-a}] \cong$

$(\mathbb{Z}/p^{n-a}\mathbb{Z})^2$. By Corollary 1.5, $V[p^{n-a}]/\overline{\Lambda} \cong \mathbb{Z}/p^{n-2a}\mathbb{Z}$. By duality the number of quotient groups of $V[p^{n-a}]$ which are cyclic of order p^{n-2a} is equal to the number of subgroups of $V[p^{n-a}]$ which are cyclic of order p^{n-2a}; since such subgroups are contained in $V[p^{n-2a}]$, their number is $C(2, p^{n-2a}) = p^{n-2a} + p^{n-2a-1}$. This establishes the claim. Finally, note that adding up the number of type a subgroups for $0 \leq a \leq \lceil \frac{n-1}{2} \rceil$ indeed gives $p^n + \ldots + p + 1 = \sigma(p^n)$. □

Exercise: Try to adapt the above methods to compute L_N in the general case.

It is now time to admit that we have been playing around a bit: with the right tool, the computation of L_N can be done in one fell swoop. That tool is Hermite normal form. Indeed, let $\Lambda \subset \mathbb{Z}^N$ be an index n sublattice. Choosing a $\mathbb{Z}$-basis $\mathfrak{v} = \{v_1, \ldots, v_N\}$ and taking $M_{\mathfrak{v}}$ the matrix with columns $v_1, \ldots, v_N$, we have $\Lambda = M_{\mathfrak{v}}\mathbb{Z}^N$ and $n = |\det M_{\mathfrak{v}}|$. Now we encounter an issue which will recur throughout these notes: we want to consider (here: to count) lattices of index n, not ordered $\mathbb{Z}$-bases of lattices of index n. In slightly fancier terms, we want to count the number of $\mathrm{GL}_N(\mathbb{Z})$-orbits on the set of integral $N \times N$ matrices with determinant n. Aha! By Theorem 1.3 every $\mathrm{GL}_N(\mathbb{Z})$-orbit contains a unique matrix in Hermite normal form, so we need only count the number of $N \times N$ matrices in Hermite normal form with determinant n.

Let us begin with the case $N = 2$: then a determinant n matrix in HNF is of the form $\left[\begin{array}{cc} a & b \\ 0 & \frac{n}{a} \end{array} \right]$ with $a > 0$ and $0 \leq b < a$. Thus a ranges over all (positive!) divisors of n and for each such divisor a we have a choices for b, and thus $\sum_{d|I} n = \sigma(n)$ HNF matrices in all. This gives a new proof of Theorem 1.12!

The case of general N is not essentially harder but only requires a little more notation. Fix N and n, and consider the set of $N \times N$ matrices in Hermite normal form of determinant N. The upper left entry can be any positive divisor d of n; the remaining $N-1$ entries on the first row are arbitrary elements of $[0, d-1)$, hence there are d^{N-1} choices. The entries in the first column below the first are all zero, and – the key point! – the lower right $N-1 \times N-1$ submatrix is also in Hermite normal form and has determinant $\frac{n}{d}$. It follows that

$$L_N(n) = \sum_{d|n} d^{N-1} L_{N-1}\left(\frac{n}{d}\right). \tag{1}$$

A student of elementary number theory will recognize the right hand side of (1) as a **Dirichlet convolution**: in general, for functions $f, g : \mathbb{Z}^+ \to \mathbb{C}$, we put

$$(f * g)(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right).$$

The Dirichlet convolution is (easily seen to be) a commutative, associative product (in fact, the set of all functions $f : \mathbb{Z}^+ \to \mathbb{C}$ endowed with pointwise sum and convolution product is an interesting commutative ring, in particular a UFD). Let us write I for the function $n \mapsto n$. Then we may rewrite (1) as

$$\forall N \geq 2,\ L_N = I^{N-1} * L_{N-1},$$

from which we immediately deduce the following result.

Theorem 1.13. *For $N \in \mathbb{Z}^+$,*

$$L_N = I^{N-1} * I^{N-2} * \cdots I \cdot I^0.$$

In other words, for all $n \in \mathbb{Z}^+$,

$$L_N(n) = \sum_{n_1,\ldots,n_N | n,\ n_1 \cdots n_N = n} n_1^{N-1} n_2^{N-2} \cdots n_{N-1}^1 n_N^0.$$

Remark 1. *We get an even prettier formula by performing a "Dirichlet transform". Given a function $f : \mathbb{Z}^+ \to \mathbb{C}$, we associate the formal Dirichlet series*

$$D_f(s) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s}.$$

*Let us define the **zeta function** of the $\mathbb{Z}$-module $\mathbb{Z}^N$ as*

$$\zeta_{\mathbb{Z}^N}(s) = D_{L_N}(s) = \sum_{n=1}^{\infty} \frac{L_N(n)}{n^s}.$$

Then Theorem 1.13 is equivalent to the identity

$$\zeta_{\mathbb{Z}^N}(s) = \zeta(s)\zeta(s-1)\cdots\zeta(s-(N-1)),$$

where $\zeta(s)$ is the usual Riemann zeta function.

Unfortunately it is not always trivial to wring from an "exact formula" a useful asymptotic estimate. Here we can give a useful upper bound.

Corollary 1.14. *Fix $N \in \mathbb{Z}^+$. Then the number of sublattices of $\mathbb{Z}^N$ of index at most T is at most T^N.*

Proof. It will be convenient to allow T to be any positive *real* number and define $\mathbb{L}_N(T) = \sum_{n=1}^{\lfloor T \rfloor} L_N(n)$; thus we will show that $\mathbb{L}_N(T) \leq T^N$ for all $T > 0$, and we do so by induction on N, the case $N = 1$ being clear, since $\mathbb{L}_1(T) = \lfloor T \rfloor \leq T$. So let $N \geq 2$ and assume that $\mathbb{L}_{N-1}(T) \leq T^{N-1}$. Then

$$\mathbb{L}_N(T) = \sum_{n_1,\ldots,n_N > 0,\ n_1 \cdots n_N \leq T} n_1^{N-1} n_2^{N-2} \cdots n_{N-1}^1 n_N^0$$

$$= \sum_{n_1=1}^{\lfloor T \rfloor} \sum_{n_2,\ldots,n_N > 0,\ n_2 \cdots n_N \leq T/n_1} n_2^{N-2} \cdots n_{N-1}^1 n_N^0 = \sum_{n_1=1}^{\lfloor T \rfloor} n_1^{N-1} \mathbb{L}_{N-1}(T/n_1)$$

$$\overset{\text{IH}}{\leq} \sum_{n_1=1}^{\lfloor T \rfloor} n_1^{N-1} \left(\frac{T}{n_1}\right)^{N-1} = \sum_{n_1=1}^{\lfloor T \rfloor} T^{N-1} = \lfloor T \rfloor T^{N-1} \leq T^N.$$

□

In fact:

Corollary 1.15. *Fix $N \in \mathbb{Z}^+$. Then as a function of T,*

$$\sum_{n=1}^{T} L_N(n) \sim \left(\frac{\zeta(2)\zeta(3)\cdots\zeta(N)}{N}\right) T^N.$$

Proof. We refer the interested reader to [GG06, Lemma 1.1]. □

1.3.3. *A characterization of lattice subbases.*

Lemma 1.16. *(Hermite's Lemma) Let R be a PID with fraction field K, let $0 \neq v = (a_1, \ldots, a_n) \in R^n$, and let be any generator of the ideal $\langle a_1, \ldots, a_n \rangle$.*
a) There exists $A \in M_n(R)$ with first column v and $\det A = d$.
b) In particular, the unimodular group $\mathrm{GL}_n(R)$ acts transitively on primitive vectors in R^n, i.e., vectors $v = (a_1, \ldots, a_n)$ with $\langle a_1, \ldots, a_n \rangle = R$.

Proof. a) We go by induction on n, the case $n = 1$ being trivial. The case $n = 2$ is easy, but let's do it: by definition of d, there exist $b_1, b_2 \in R$ such that $a_1 b_2 - a_2 b_1 = d$, and then the matrix

$$A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$$

has first column $v = (a_1, a_2)$ and determinant d.
Now assume that $n \geq 3$ and that the result holds in R^{n-1}. Thus there exists $A' \in M_{n-1}(R)$ with first column $(a_1, \ldots, a_{n-1})$ and determinant d' any prescribed generator of the ideal $\langle a_1, \ldots, a_{n-1} \rangle$. Since $dR = \langle a_1, \ldots, a_{n-1}, a_n \rangle$, there are $x, y \in R$ such that $d'x - a_n y = d$. Now we consider the following matrix $A \in M_n(R)$: its upper left $(n-1) \times (n-1)$ corner will be the matrix A'; its nth row will be $(a_n, 0, \ldots, 0, x)$, and its nth column will be $(\frac{a_1 y}{d'}, \ldots, \frac{a_{n-1} y}{d'}, x)^T$. Note that A has first column $(a_1, \ldots, a_n)^T$, so it remains to show that $\det A = d$.

Let $A_{n,1}$ be the minor obtained by crossing out the nth row and 1st column of A. Then Laplace expansion along the nth row of A gives

$$\det A = (-1)^{n-1} a_n \det A_{n-1} + d'x.$$

Moreover, the matrix $d' A_{n-1}$ is obtained from A' by multiplying columns 2 through $n-1$ of A' by d', then cyclically permuting the columns, and finally multiplying the last column by y. Thus

$$(d')^{n-1} \det A_{n-1} = \det(d' A_{n-1}) = (d')^{n-2} (-1)^{n-2} y \det A' = (d')^{n-1} (-1)^{n-2} y,$$

so

$$\det A_{n-1} = (-1)^{n-2} y,$$

and thus

$$\det A = (-1)^{n-1} a_n (-1)^{n-2} y + d'x = d'x - a_n y = d.$$

b) For $A \in M_n(R)$, $A \in \mathrm{GL}_n(R)$ iff $\langle \det A \rangle = R$, so this follows from part a). □

The above proof of Lemma 1.16a) is a very classical one. (I don't know whether it was Hermite's proof, but it uses only tools that he would have had.) On the other hand, if we think about Lemma 1.16 from the perspective of module theory, one can give a much simpler argument.[2] We develop it in the following exercises.

Exercise: Let R be an integral domain with fraction field K. A vector $v = (x_1, \ldots, x_n) \in R^n$ is **primitive** if $\langle x_1, \ldots, x_n \rangle = R$.
a) Let $b_1, \ldots, b_n$ be a basis for R^n. Show that each b_i is a primitive vector.
b) Show that for $v \in (R^n)^\bullet$, the following are equivalent:
(i) v is a primitive vector.
(ii) $\langle v \rangle_K \cap R^n = \langle v \rangle$.
(iii) The R-module $R^n / \langle v \rangle$ is torsionfree.

[2] I learned this simpler approach from Martin Brandenburg.

c) Let R be a PID and $v \in R^n$ be a primitive vector. Use part b) and the short exact sequence

$$0 \to \langle v \rangle \to R^n \to R^n/\langle v \rangle \to 0$$

to show that there is $M \in \mathrm{GL}_n(R)$ with $M(e_1) = v$.
d) Deduce part a) of Hermite's Lemma from part c).

1.4. **The Classification of Vector Groups.**

The aim of this section is to prove that every closed subgroup of $\mathbb{R}^N$ is isomorphic to $\mathbb{R}^d \times \mathbb{Z}^{r-d}$, where $r = \mathfrak{r}(G)$ and $0 \leq d \leq N$. In fact, our first order of business is to define the quantity d for any vector group, i.e., a not necessarily closed subgroup of $\mathbb{R}^N$.

Let G be a subgroup of $\mathbb{R}^N$. For $\epsilon > 0$, we define $d(\epsilon)$ to be the maximal cardinality of an $\mathbb{R}$-linearly independent subset of $G \cap B_0(\epsilon)$, and we put

$$d = d(G) = \inf_{\epsilon > 0} d(\epsilon).$$

Since $d(\epsilon)$ is a weakly increasing function of ϵ taking values in the finite set $\{0, \ldots, N\}$, for sufficiently small values of ϵ it is constant, and this constant value is $d(G)$, the **dimension** of G.

Exercise: Show that G is discrete iff $d(G) = 0$.

To any d-dimensional vector group G we attach a d-dimensional linear subspace $W = W(G)$ as follows: choose $\epsilon_0 > 0$ sufficiently small such that $d = d(\epsilon_0)$. For any $0 < \epsilon < \epsilon_0$, choose $\mathbb{R}$-linearly independent vectors $v_1, \ldots, v_d$ and put $W(\epsilon) = \langle v_1, \ldots, v_d \rangle$. Each $W(\epsilon)$ is independent of the chosen vectors, for otherwise we would get more than s linearly independent vectors of length less than ϵ_0. By the same reasoning, for all $\epsilon_2 < \epsilon_1 < \epsilon_0$, $W(\epsilon_1) = W(\epsilon_2)$, and thus we have defined a subspace W depending only on G.

Let us put $G_c = G \cap W(G)$.

Lemma 1.17. *G_c is dense in $W(G)$.*

Theorem 1.18. *Let G be a vector group with $\mathfrak{r}(G) = r$ and dimension d. Then there is a discrete subgroup G_d of G of rank $r - d$ such that*

$$G = G_c \oplus G_d.$$

Corollary 1.19. *Let $G \subset \mathbb{R}^N$ be a closed vector group. Then $G_c = W(G)$, so*

$$\overline{G} = W \oplus G' \cong \mathbb{R}^d \oplus \mathbb{Z}^{r-d}.$$

Proof. By Lemma 1.17 G_c is dense in $W(G)$. But since G is closed in $\mathbb{R}^N$, so is $G_c = G \cap W(G)$, and thus $G_c = W(G)$. The rest follows from Theorem 1.18. □

1.5. **The space of all lattices.**

Let $\mathcal{L}_N$ be the set of all lattices in $\mathbb{R}^N$. The linear action of $\mathrm{GL}_N(\mathbb{R})$ on $\mathbb{R}^N$ induces an action of $\mathrm{GL}_N(\mathbb{R})$ on $\mathcal{L}_N$: namely, for $g \in \mathrm{GL}_N(\mathbb{R})$ and $\Lambda \in \mathcal{L}_N$, we put

$$g\Lambda = \{gx \mid x \in \Lambda\}.$$

To see that $g\Lambda$ is again a lattice, choose a basis $v_1, \ldots, v_N$ for Λ. Then $gv_1, \ldots, gv_N$ is an $\mathbb{R}$-basis for $\mathbb{R}^N$ and a $\mathbb{Z}$-basis for $g\Lambda$, so $g\Lambda$ is a lattice. Further, since every lattice has a basis and $\mathrm{GL}_N(\mathbb{R})$ acts transitively on bases of $\mathbb{R}^N$, it follows that $\mathrm{GL}_N(\mathbb{R})$ acts transitively on $\mathcal{L}_N$. However, unlike the action on bases, since a lattice has many bases, the action on $\mathrm{GL}_N(\mathbb{R})$ is not simply transitive. The stabilizer of the standard integer lattice $\mathbb{Z}^N$ in $\mathrm{GL}_N(\mathbb{R})$ is, essentially by definition, the discrete subgroup $\mathrm{GL}_N(\mathbb{Z})$, i.e., the subgroup of invertible matrices with integer entries and with inverse having integer entries, or equivalently the subgroup of matrices having integer entries and determinant ± 1. The orbit stabilizer theorem gives an isomorphism of $\mathrm{GL}_N(\mathbb{R})$-sets

$$\mathcal{L}_N \cong \mathrm{GL}_N(\mathbb{R})/\mathrm{GL}_N(\mathbb{Z}).$$

By some deep theorems in differential geometry, $\mathrm{GL}_N(\mathbb{R})/\mathrm{GL}_N(\mathbb{Z})$ can be naturally endowed with the structure of a smooth (even real analytic) manifold, whose underlying topology is simply the quotient topology. Thus we may view $\mathcal{L}_N$ as having this structure.

Example 1.5: When $N = 1$, $\mathrm{GL}_N(\mathbb{R})/\mathrm{GL}_N(\mathbb{Z}) \cong \mathbb{R}^{>0}$: this corresponds to the fact that a lattice in $\mathbb{R}$ has a unique positive real number generator. Notice that this space is non-compact in "two different directions". This will later be made precise, and necessary and sufficient conditions for a subset of $\mathcal{L}_N$ to be compact will be given: **Mahler's Compactness Theorem**.

2. The Lattice Point Enumerator

2.1. **Introduction.**

Consider the following very classical problem: how many lattice points lie on or inside the circle $x^2 + y^2 = r^2$? Equivalently, for how many pairs $(x, y) \in \mathbb{Z}^2$ do we have $x^2 + y^2 \leq r^2$? Let $L(r)$ denote the number of such pairs.

Upon gathering a bit of data, it becomes apparent that $L(r)$ grows quadratically with r, which leads to consideration of $\frac{L(r)}{r^2}$. Now:

$$L(10)/10^2 = 3.17.$$
$$L(100)/100^2 = 3.1417.$$
$$L(1000)/1000^2 = 3.141549.$$
$$L(10^4)/10^8 = 3.14159053.$$

The pattern is pretty clear!

Theorem 2.1. *As $r \to \infty$, we have $L(r) \sim \pi r^2$. Explicitly,*

$$\lim_{r \to \infty} \frac{L(r)}{\pi r^2} = 1.$$

Once stated, this result is quite plausible geometrically: suppose that you have to tile an enormous circular bathroom with square tiles of side length 1 cm. The total number of tiles required is going to be very close to the area of the floor in square centimeters. Indeed, starting somewhere in the middle you can do the vast majority of the job without even worrying about the shape of the floor. Only when you come within 1 cm of the boundary do you have to worry about pieces of tiles and

so forth. But the number of tiles required to cover the boundary is something like a constant times the perimeter of the region in centimeters – so something like $C\pi r$ – whereas the number of tiles in the interior is close to πr^2. Thus the contribution to the boundary is neglible: precisely, when divided by r^2, b it approaches 0 as $r \to \infty$.

I myself find this heuristic convincing but not quite rigorous. More precisely, I believe it for a circular region and become more concerned as the boundary of the region becomes more irregularly shaped, but the heuristic doesn't single out exactly what nice properties of the circle are being used. Moreover the "error" bound is fuzzy: it would be useful to know an explicit value of C.

2.2. **Gauss's Solution to the Gauss Circle Problem.**

The first proof of Theorem 2.1 that we will present was given by Gauss in 1837. In fact he proves a stronger result. Namely, we define the **error**

$$E(r) = |L(r) - \pi r^2|,$$

so that Theorem 2.1 is equivalent to the statement

$$E(r) = o(r^2),$$

or to spell out the "little oh notation",

$$\lim_{r\to\infty} \frac{E(r)}{r^2} = 0.$$

Theorem 2.2. *(Gauss) For all $r \geq 7$, $E(r) \leq 10r$.*

Proof. Let $P = (x, y) \in \mathbb{Z}^2$ be such that $x^2 + y^2 \leq r^2$. To P we associate the square $S(P) = [x, x + 1] \times [y, y + 1]$, i.e., the unit square in the plane which has P as its lower left corner. Note that the diameter of $S(P)$ – i.e., the greatest distance between any two points of $S(P)$ – is $\sqrt{2}$. So, while P lies within the circle of radius r, $S(P)$ may not, but it certainly lies within the circle of radius $r + \sqrt{2}$. It follows that the total area of all the squares $S(P)$ – which is nothing else than the number $L(r)$ of lattice points – is at most the area of the circle of radius $r + \sqrt{2}$, i.e.,

$$L(r) \leq \pi(r + \sqrt{2})^2 = \pi r^2 + 2\sqrt{2}\pi r + 2.$$

A similar argument gives a lower bound for $L(r)$. Namely, if (x, y) is any point with distance from the origin at most $r - \sqrt{2}$, then the entire square $(\lfloor x \rfloor, \lfloor x + 1 \rfloor) \times (\lfloor y \rfloor, \lfloor y + 1 \rfloor)$ lies within the circle of radius r. Thus the union of all the unit squares $S(P)$ attached to lattice points on or inside $x^2 + y^2 = r$ covers the circle of radius $r - \sqrt{2}$, giving

$$L(r) \geq \pi(r - \sqrt{2})^2 = \pi r^2 - 2\sqrt{2}\pi r + 2.$$

Thus

$$E(r) = |L(r) - \pi r^2| \leq 2\pi + 2\sqrt{2}\pi r \leq 7 + 9r \leq 10r,$$

the last inequality holding for all $r \geq 7$. □

2.3. **A second solution to the Gauss Circle Problem.**

Here is a second, quite different, proof of Theorem 2.1.

The first step is to notice that instead of counting lattice points in an expanding sequence of closed disks, it is equivalent to fix the plane region once and for all – here, the unit disk $D : x^2+y^2 \leq 1$ – and consider the number of points $(x, y) \in \mathbb{Q}^2$ with $rx, ry \in \mathbb{Z}$. That is, instead of dividing the plane into squares of side length one, we divide it into squares of side length $\frac{1}{r}$. If we now count these "$\frac{1}{r}$-lattice points" inside D, a moment's thought shows that this number is precisely $L(r)$.

Now what sort of thing is an area? In calculus we learn that areas are associated to integrals. Here we wish to consider the area of the unit disk as a **double integral** over the square $[-1, 1]^2$. In order to do this, we need to integrate the **characteristic function** $\mathbf{1}_D$ of the unit disk: that is, $\mathbf{1}_D(x)$ is 1 if $x \in D$ and otherwise. Now the division of the square $[-1, 1]^2$ into $4r^2$ subsquares of side length $\frac{1}{r}$ is exactly the sort of sequence of partitions that we need to define a Riemann sum: that is, the maximum diameter of a subrectangle in the partition is $\frac{\sqrt{2}}{r}$, which tends to 0 as $r \to \infty$. Therefore if we choose any point $P_{i,j}^*$ in each subsquare, then

$$\Sigma_r := \frac{1}{r^2} \sum_{i,j} \mathbf{1}_D(P_{i,j}^*)$$

is a sequence of Riemann sums for $\mathbf{1}_D$, and thus

$$\lim_{r\to\infty} \Sigma_r = \int_{[-1,1]^2} \mathbf{1}_D = \text{Area}(D) = \pi.$$

But we observe that Σ_r is very close to the quantity $L(r)$. Namely, if we take each sample point to be the lower left corner of corner of the corresponding square, then $r^2\Sigma_r = L(r) - 2$, because every such sample point is a lattice point (which gets multiplied by 1 iff the point lies inside the unit circle) and the converse is true except that the points $(1, 0)$ and $(0, 1)$ are not chosen as sample points. So

$$\lim_{r\to\infty} \frac{L(r)}{r^2} = \lim_{r\to\infty} \frac{L(r) - 2 + 2}{r^2} = \lim_{r\to\infty} \Sigma_r + 0 = \pi.$$

2.4. **Introducing the Lattice Point Enumerator.**

One may well wonder why we have bothered with the second proof of Theorem 2.1 since the first proof is more elementary and gives a sharper result. The answer is that the second proof is amenable to a significant generalization. Indeed, consider any bounded subset $\Omega \subset \mathbb{R}^N$, and for $r \in \mathbb{R}^{>0}$ consider the r-**dilate** of Ω:

$$r\Omega = \{rP = (rx_1, \ldots, rx_N) \mid P = (x_1, \ldots, x_n) \in \Omega\}.$$

We define the **lattice point enumerator**

$$L_\Omega(r) = \#(r\Omega \cap \mathbb{Z}^N),$$

which counts the number of standard lattice points lying in the rth dilate of Ω. Now we wish to generalize Theorem 2.1 by showing that as r approaches ∞, $L_\Omega(r)$ is asymptotic to the (N-dimensional) volume of Ω times r^N.

However, we certainly need *some additional hypothesis* on Ω for this to be true.

Example: Let $[0,1]^N_{\mathbb{Q}}$ be the set of all points $(x_1, \ldots, x_N$ with each x_i a *rational number* in $[0,1]$, and put $\Omega = [0,1]^N \setminus [0,1]^N_{\mathbb{Q}}$. Then, for any $r \in \mathbb{Z}^+$, the lattice points in $r\Omega$ are in bijection with the $\frac{1}{r}$-lattice points in Ω, of which there are none. On the other hand, for irrational r, $L_\Omega(r)$ does grow like r^N, so overall it is not asymptotic to any constant times r^N. The set Ω is Lebesgue measurable with $m(\Omega) = 1$, so evidently requiring Lebesgue measurability is not enough.

Looking at the second proof of Theorem 2.1 we can isolate the condition on Ω that was used: the **Riemann integrability** of the characteristic function χ_D of the region D. It is a basic fact that a bounded function on a bounded domain is Riemann integrable if and only if it is continuous except on a set of measure zero. The characteristic function χ_D is discontinuous precisely along the boundary of D, so the necessary condition on D is that its boundary have measure zero. In geometric measure theory, such regions are called **Jordan measurable**.

Jordan measurability is a relatively mild condition on a region: for instance any region bounded by a piecewise smooth curve (a circle, ellipse, polygon...) is Jordan measurable. In fact a large collection of regions with fractal boundaries are Jordan measurable: for instance Theorem 2.3 applies with R a copy of the **Koch snowflake**, whose boundary is a nowhere differentiable curve.

Note that we have defined Jordan measurability and not Jordan *measure*. It is certainly possible to do so: roughly speaking, we define outer and inner Jordan measure as with Lebesgue measure but by using **finite** unions of basic regions (products of intervals), and then the most salient feature of the Jordan measurable sets are that they form an algebra of sets but not a σ-algebra. It is not yet clear to me whether it is our business to get into the finer points of Jordan measure, so I have not included it here (or really learned it myself...). In particular every Jordan measurable set is Lebesgue measurable, and we denote the Lebesuge measure of $\Omega \subset \mathbb{R}^N$ simply as $\operatorname{Vol} \Omega$.

Theorem 2.3. *Let $\Omega \subset \mathbb{R}^N$ be bounded and Jordan measurable. Then*

$$\lim_{r \to \infty} \frac{L_\Omega(r)}{r^N} = \operatorname{Vol} \Omega.$$

Proof. This is a very direct generalization of the second proof of Theorem 2.1 given in §1.3 above. Indeed, by scaling appropriately we may assume that $\Omega \subset (-1,1)^N$, and then for any $r \in \mathbb{Z}^+$, $\frac{L_\Omega(r)}{r^N}$ is *precisely* a Riemann sum for the characteristic function $\mathbf{1}_\Omega$ corresponding to the partition of $[-1,1]^N$ into subsquares of side length $\frac{1}{r}$, so the convergence of these sums to $\int \mathbf{1}_\Omega$ as $r \to \infty$ is immediate from the definition of Jordan measurability.[3] □

Exercise: Show that a bounded, Jordan measurable set has positive volume iff it has nonempty interior.

[3]One technical remark: when we say the limit exists as $r \to \infty$, we really mean that r is allowed to take on all positive real number values. In order to divide $[-1,1]^N$ *exactly* into subsquares of side length $\frac{1}{r}$ we clearly need $r \in \mathbb{Z}^+$. However, it should be rather clear that the argument can be adapted to the case of arbitrary positive r at the cost of making it a little less clean.

When Ω has nonempty interior – the case that we will be mostly interested in in what follows – an equivalent statement of Theorem 2.3 is

$$L_\Omega(r) \sim (\operatorname{Vol}\Omega) r^N \tag{2}$$

as $r \to \infty$. (When $\operatorname{Vol}\Omega = 0$, (2) asserts that for sufficiently large r, $r\Omega$ contains no lattice points, which of course need not be true. I thank David Krumm for pointing out that Theorem 2.3 should not be formulated this way when $\operatorname{Vol}\Omega = 0$, although I didn't immediately understand what he was getting at.)

2.5. **Error Bounds on the Lattice Enumerator.**

Suppose $\Omega \subset \mathbb{R}^N$ is a bounded Jordan measurable set of positive volume (equivalently, nonempty interior). By a simple Riemann integration argument we showed $L_\Omega(r) \sim (\operatorname{Vol}\Omega) r^N$ as $r \to \infty$. Again though, when Ω is the closed unit disk in $\mathbb{R}^2$ Gauss's classical argument did better than this: we got not only an asymptotic formula for the lattice point enumerator but an **explicit upper bound** for the error function

$$E(r) = |L_\Omega(r) - (\operatorname{Vol}\Omega) r^N|.$$

So it is a natural problem – perhaps the **fundamental problem** in this area – to give sharp bounds on $E(r)$ for various Jordan measurable regions Ω.

A little reflection shows that what we want to say about the error will depend quite a lot on what sort of set the boundary $\partial\Omega$ is. For instance, one has the following generalization of Gauss's argument.

Theorem 2.4. *Suppose that $\Omega \subset \mathbb{R}^N$ is bounded, Jordan measurable with nonempty interior and that $\partial\Omega$ is piecewise C^1. Then*

$$E(r) = O(r^{N-1}).$$

I find Theorem 2.4 to be "geometrically obvious", and there is even a published paper of mine in which I simply assert it without any proof or reference. Nevertheless it would be a nice exercise for someone to write down a careful argument.

Example 2.5. *Let $\Omega = [-1,1]^2 \subset \mathbb{R}^2$. Then for all $r \in \mathbb{Z}^+$, $r\Omega = [-r,r]^2$, and this square consists of $2r+1$ rows of $2r+1$ lattice points, so*

$$L_\Omega(r) = (2r+1)^2 = 4r^2 + 4r + 1$$

and thus

$$E_\Omega(r) = 4r + 1.$$

That is, the error term actually is *as large as a constant times r^{N-1} in this case. The coefficient 4 of r^2 is indeed the 2-dimensional volume of $\Omega = [-1,1]^2$. The coefficient 4 of r is in fact half the 1-dimensional volume of the boundary $\partial\Omega$ (in this case the perimeter of the square, which is 8). The constant coefficient 1 is in fact the **Euler characteristic** of Ω.*

Exercise: Extend Example 2.5 to $[-1,1]^N \subset \mathbb{R}^N$.

Remark 2. *In some sense $[-1,1]^2$ is the "worst placement" of the unit square: it is positioned so as to pick up as many lattice points on the boundary as possible. If*

we were to ***rotate*** *the square about the origin by some generic*[4] *angle, we have a right to expect a smaller error term.*

To say more about the error term I now want to consider separately three different cases.

2.6. **When $\partial\Omega$ is smooth and positively curved.**

The first person to make qualitative progress on Gauss's bound $E(r) = O(r)$ for Ω the unit disk in $\mathbb{R}^2$ was W. Sierpinski, in a prize essay he submitted while an undergraduate (!) student at the Univesity of Warsaw. He showed:

Theorem 2.6. *(Sierpinski, 1906) Let Ω be the closed unit disk in $\mathbb{R}^2$. Then*

$$E_\Omega(r) = O(r^{\frac{2}{3}}).$$

The next important result was a lower bound on the error.

Theorem 2.7. *(Hardy, Landau 1916) Let Ω be the closed unit disk in $\mathbb{R}^2$. There is* no constant *C such that $E_\Omega(r) \leq Cr^{\frac{1}{2}}$.*

The standard conjecture is that the Hardy-Landau lower bound is essentially sharp.

Conjecture 2.8. *(Gauss Circle Problem) Let Ω be the closed unit disk in $\mathbb{R}^2$. Then for every $\epsilon > 0$, there exists $C_\epsilon > 0$ (i.e., a "constant depending on ϵ") such that*

$$E_\Omega(r) \leq C_\epsilon r^{\frac{1}{2}+\epsilon}.$$

So far as I know, the best known upper bound is the following one.

Theorem 2.9. *(Huxley* [Hu00]*) Let Ω be the closed unit disk in $\mathbb{R}^2$. Then*

$$E_\Omega(r) = O(r^{\frac{131}{208}}) = O(r^{0.6298\cdots}).$$

Despite the fact that $\frac{131}{208}$ is much closer to $\frac{2}{3}$ than to $\frac{1}{2}$, in between Sierpinski and Huxley come many other distinguished mathematicians. In other words, Conjecture 2.8 is extremely difficult. In particular it is beyond the scope of our research group to work on, and I mention it just for culture.

Because the disk is rotationally symmetric, it is reasonable to expect that its error function is especially small. But one wants to prove similar, if more modest, upper bounds for regions Ω which have smooth, positively curved boundary, e.g. ellipsoids. For this one has the following result.

Theorem 2.10. *(van der Corput, Hlawka) Suppose that $\partial\Omega$ is sufficiently smooth with everywhere positive Gaussian curvature. Then*

$$E_\Omega(r) = O(r^{n\left(\frac{n-1}{n+1}\right)}).$$

Taking $N = 2$ in Theorem 2.10 gives an exponent of $\frac{2}{3}$ and hence a generalization of Theorem 2.6. The case $N = 2$ was established by van der Corput in his 1919 thesis. The higher dimensional case is due to Hlawka.

[4]Using this word probably reveals my loyalties to algebraic geometry rather than geometric measure theory. Better would be: *random*.

2.7. When $\partial\Omega$ is a fractal set.

Given that the science of counting lattice points in regions with very nice boundaries is probably too advanced for us to jump in midstream and make any kind of meaningful contribution, it is tempting to switch to the *other* extreme: recall that in Theorem 2.3 we are allowed to take any bounded, Jordan measurable set Ω, i.e., any bounded set whose boundary has Lebesgue measure zero. In particular, $\partial\Omega$ can be a **fractal set** of some fractal dimension $N-1 < \alpha < N$. For instance, we could count lattice points in dilates of the **Koch snowflake**.

It is natural to expect an upper bound on the error in terms of the fractal dimension of $\partial\Omega$. This is attained in relatively recent work of L. Colzani.

Theorem 2.11. *(Colzani* [Col97]*) Let $\Omega \subset \mathbb{R}^N$ be a bounded Jordan measurable set with boundary of fractal dimension α. Then*

$$E_\Omega(r) = O(r^\alpha).$$

Remark 3. *a) We are being deliberately vague with the term "fractal dimension", since there are many different definitions of fractal dimension, morally the same but differing in their technical details. Colzani proves the result for a specific notion of fractal dimension adapted to his purpose, which seems reasonable. Whether the result holds for, say, the Hausdorff dimension is unknown to me.*
b) If for $\Omega \subset \mathbb{R}^N$, $\partial\Omega$ is piecewise C^1 then it will have fractal dimension $N-1$. Thus Colzani's Theorem should in particular recover Theorem 2.4, and this is worth checking up on.

So far as I know, the question of a converse to Theorem 2.11 remains open.

Problem 1. *For each $N \in \mathbb{Z}^+$ and $\alpha \in (N-1, N)$, find a bounded Jordan measurable subset $\Omega \subset \mathbb{R}^N$ such that $E_\Omega(r)$ is* not *bounded by a constant times r^α.*

I haven't seriously thought about Problem 1, so as far as I know it could be rather simple to prove (or not, of course). It's certainly worth a try.

2.8. When Ω is a polytope.

We will devote the next section to a study of this case.

3. The Ehrhart (Quasi-)Polynomial

3.1. Basic Terminology.

For a subset $S \subset \mathbb{R}^N$, the **affine hull** of S is the least affine linear subspace containing S (that is, the intersection of all affine subspaces of $\mathbb{R}^N$ containing S). If S is convex, then the **dimension of S** is the dimension of the affine hull of S. (This is a reasonable definition: any convex subset, viewed as a subspace of its affine hull, has nonempty interior, and thus e.g. the Hausdorff dimension of S is equal to the dimension of its affine hull.) We say that S is **full dimensional** if its affine hull is $\mathbb{R}^N$.

A **convex polytope** in Euclidean space $\mathbb{R}^N$ is the convex hull of a finite subset of points. A convex polytope is **integral** if it is the convex hull of a finite

subset of $\mathbb{Z}^N$ and **rational** if it is the convex hull of a finite subset of $\mathbb{Q}^N$.

Let $L \in \mathbb{Z}^+$ and $d \in \mathbb{N}$. A function $f : \mathbb{Z}^+ \to \mathbb{C}$ is a **quasi-polynomial function of period L and degree d** if it can be expressed in the form

$$f(r) = c_d(r)r^d + c_{d-1}r^{d-1} + \ldots + c_1(r)r + c_0(r)$$

for L-periodic functions $c_0, \ldots, c_d : \mathbb{Z}^+ \to \mathbb{C}$ with $c_d(r)$ not identically zero. Just to be sure, we say $c_i : \mathbb{Z}^+ \to \mathbb{C}$ is L- periodic if $x \equiv y \pmod{L} \implies c_i(x) = c_i(y)$. Thus a quasi-polynomial function is a function which is given as a possibly different polynomial on each residue class modulo L, and such that the largest degree of these polynomials is equal to d. Note also that saying that a function has period L does not preclude the possibility that it may also have period some proper divisor of L. Finding the least period of a quasi-polynomial can be an interesting problem.

Theorem 3.1. *(Ehrhart) Let $\Omega \subset \mathbb{R}^N$ be a polytope with nonempty interior.*
a) If Ω is an integral polytope, there is a polynomial $P(t) \in \mathbb{R}[t]$ such that for all $r \in \mathbb{Z}^+$, $L_\Omega(r) = P(r)$.
b) If Ω is a rational polytope and $L \in \mathbb{Z}^+$ is a common denominator for the coordinates of the vertices of Ω, then there is a quasi-polynomial $f : \mathbb{Z}^+ \to \mathbb{R}$ of period L and degree N such that $f(r) = L_\Omega(r)$ for all r.

Exercise: Let Ω be an integral polytope with Ehrhart polynomial $P(t) = c_d t^d + \ldots + c_1 t + c_0$ of degree d. Show that for all $0 \leq k \leq d$, $d!c_k \in \mathbb{Z}$. (Hint: this has nothing to do with polytopes or geometry. In fact it holds for any polynomial which take integer values at all positive integers.)

The polynomial in part a) is the **Ehrhart polynomial** and the quasi-polynomial in part b) is the **Ehrhart quasi-polynomial**. Both are worthy objects of study.

Just to name one specific problem, it is interesting to look at how many distinct polynomials comprise the Ehrhart quasi-polynomial: *a priori* we may get a different polynomial function for each residue class modulo the least common multiple of the denominators of the vertices of Ω. But it follows from Theorem 2.1 that the leading coefficient of every quasi-polynomial is $\operatorname{Vol}\Omega$. Roughly speaking the coefficient $c_k(r)$ of r^k "depends more and more strongly on r" as k decreases.

Problem 2. *Let $\Omega \subset \mathbb{R}^N$ be a rational polytope.*
a) What is the period of the Ehrhart quasi-polynomial? Of each coefficient?
b) How many distinct polynomial functions comprise the Ehrhart quasi-polynomial?

These questions are interesting already for very simple polytopes. In [AC05], Gil Alon and I looked at the case of the simplices $ax + by + cz = 1$, $x, y, z \geq 0$. (Note that this is a 2-dimensional simplex – i.e., a triangle – living inside a hyperplane in $\mathbb{R}^3$. Thus it is not a "full dimensional polytope" in $\mathbb{R}^3$ so does not literally fit into the setup of Theorem 3.1, but the results adapt immediately to such things.) We found that taking the coefficients $a, b, c \in \mathbb{Z}$ to be *not pairwise coprime* has interesting effects on the Ehrhart coefficients.

4. Convex Sets, Star Bodies and Distance Functions

4.1. **Centers and central symmetry.**

Let $\Omega \subset \mathbb{R}^N$. A point C is a **center** for Ω if for all $P \in \Omega$, the reflection of P through C is also in Ω.

Exercise: A bounded subset of $\mathbb{R}^N$ has at most one center.

Thus, by choosing coordinates appropriately we may assume that a bounded set with a center has the origin as the center: we say that a subset Ω of $\mathbb{R}^N$ is **centrally symmetric** if the origin is a center, or in other words if $P \in \Omega \implies -P \in \Omega$.

4.2. **Convex Subsets of Euclidean Space.**

A subset S of $\mathbb{R}^N$ is **convex** if $x, y \in S \implies \lambda x + (1-\lambda)y \in S$ for all $0 \leq \lambda \leq 1$.

Exercise: Is the empty set convex?

Exercise:
a) Show: if $\{S_i\}_{i \in I}$ is any family of convex subsets of $\mathbb{R}^N$, then $\bigcap_{i \in I} S_i$ is convex.
b) Show that the union of two convex subsets of $\mathbb{R}^N$ need not be convex.
c) Let $S_1 \subset S_2 \subset \ldots \subset S_n \subset \ldots \subset \mathbb{R}^N$ all be convex. Show $\bigcup_{n=1}^{\infty} S_n$ is convex.

Exercise: Show that a nonempty convex set is connected and simply connected.

Exercise: Let $S_1 \subset \mathbb{R}^M$ and $S_2 \subset \mathbb{R}^N$ be two convex sets. Show that the Cartesian product $S_1 \times S_2 \subset \mathbb{R}^{M+N}$ is convex.

Exercise: Let S be any subset of $\mathbb{R}^N$. Define $\operatorname{Conv} S$ to be the intersection of all convex subsets Ω of $\mathbb{R}^N$ which contain S. Show that $\operatorname{Conv} S$ is the unique minimal convex subset containing S. It is often called the **convex hull** of S.

Exercise: A set $S \subset \mathbb{R}^N$ has the **midpoint property** if for all $x, y \in S$, $\frac{x+y}{2} \in S$.
a) Show that every convex set has the midpoint property.
b) Give an example of a non-convex subset of $\mathbb{R}^N$ with the midpoint property.
c) Dis/prove: an open subset $S \subset \mathbb{R}^N$ with the midpoint property must be convex.
d) Dis/prove: a closed subset $S \subset \mathbb{R}^N$ with the midpoint property must be convex.

If $x_1, \ldots, x_m$ is a finite set of vectors in $\mathbb{R}^N$, a **convex combination** of these vectors is a linear combination

$$\lambda_1 x_1 + \ldots + \lambda_m x_m$$

with $\lambda_i \in \mathbb{R}$ and satisfying the additional conditions

$$\lambda_1, \ldots, \lambda_m \geq 0, \ \lambda_1 + \ldots + \lambda_m = 1.$$

More generally, if S is any subset of $\mathbb{R}^N$, then a convex combination of elements of S is a convex combination of some finite subset of S.

Proposition 4.1. *A subset S of $\mathbb{R}^N$ is convex iff every convex combination of vectors in S is again a vector in S.*

Exercise: Prove Proposition 4.1.

Proposition 4.2. *Let S be a subset of $\mathbb{R}^N$. Then the set of all convex combinations of elements of S is* $\operatorname{Conv} S$*, the convex hull of S.*

Exercise: Prove Proposition 4.2.

For a subset $\Omega \subset \mathbb{R}^N$, we denote by $\partial\Omega$ the boundary of Ω, i.e., the closure of Ω intersected with the closure of $\mathbb{R}^N \setminus \Omega$.

Proposition 4.3. *Let K be a convex subset of $\mathbb{R}^N$.*
*a) We have $K^\circ = (\overline{K})^\circ$. (In particular an open convex set is **regular-open**.)*
*b) We have $\overline{K} = \overline{K^\circ}$. (In particular a closed convex set is **regular-closed**.)*

Proof. ... □

Proposition 4.4. *Let K be a convex subset of $\mathbb{R}^N$, and let S be a set with $K^\circ \subset S \subset \overline{K}$. Then S is convex. In particular K° and $\overline{K}$ are convex.*

Proof. ... □

To consolidate the preceding two results: a convex set need not be open or closed, but every convex set is obtained by starting with an open convex set and adding in an (arbitrary) subset of its boundary and also by starting with an closed convex set C and removing an (arbitrary) subset of its boundary. Thus it is natural to restrict attention to convex sets which are either open or closed. Moreover, there is a **duality** between open and closed convex sets: no information is lost in passing from a closed convex set to its interior or from an open convex subset to its closure. (This is useful for instance when comparing various people's definitions of "convex body": some require it to be open and some to be closed, but it certainly doesn't matter.) The use of the term "duality" for this simple observation may seem pretentious, but when we discuss the Ehrhart polynomial of integral polytope the justification will become clear!

Every convex subset of $\mathbb{R}^N$ is Lebesgue measurable. This plausible (and true!) result will be taken for granted for now. Later on we will discuss a significantly stronger result: every bounded convex subset of $\mathbb{R}^N$ is **Jordan measurable**. As is typical in this subject, we denote the N-dimensional Lebesgue measure of Ω by $\operatorname{Vol}\Omega$. (Note: "Ω is Lebesgue measurable" allows the possibility that $\operatorname{Vol}\Omega = \infty$!)

Theorem 4.5. *Let $\Omega \subset \mathbb{R}^N$ be convex.*
a) The following are equivalent:
(i) Ω is "flat": i.e., Ω is contained in some hyperplane in $\mathbb{R}^N$.
(ii) $\operatorname{Vol}\Omega = 0$.
(iii) Ω has empty interior.
b) If $0 < \operatorname{Vol}\Omega < \infty$, then Ω is bounded.

Proof. a) (i) $\implies$ (ii) $\implies$ (iii) is immediate for *all* subsets Ω of $\mathbb{R}^N$.
(iii) $\implies$ (i): we show the contrapositive. Suppose that Ω does not lie in any hyperplane in $\mathbb{R}^N$. Then there are $x_1, \ldots, x_{N+1} \in \Omega$ not lying in any hyperplane. Their convex hull is an N-dimensional simplex, which has nonempty interior.
b) By part a), since $\operatorname{Vol}\Omega > 0$, Ω has nonempty interior. Thus by translating Ω we may assume that 0 is an interior point of Ω and thus that there exists $\epsilon > 0$ such

that for all $1 \leq i \leq N$, $\epsilon e_i \in \Omega$. We CLAIM that for any $x \in \Omega$,

$$||x||_\infty = \max_{1 \leq i \leq N} |x_i| \leq \frac{(N!) \operatorname{Vol} \Omega}{\epsilon^{N-1}},$$

which certainly suffices to show that Ω is bounded.
PROOF OF CLAIM Indeed, if $x_i \neq 0$, then the simplex with vertices 0, x, $\{\epsilon e_j\}_{j \neq i}$ is contained in Ω, and this simplex has volume $\frac{\epsilon^{N-1}|x_i|}{N!} \leq \operatorname{Vol} \Omega$. □

4.3. **Star Bodies.**

Let $\Omega \subset \mathbb{R}^N$ be a subset. A point $P \in \Omega$ is a **star point** if for all $P' \in \Omega$ the entire line segment $\overline{PP'} = \{(1-\lambda)P + \lambda P' \mid 0 \leq \lambda \leq 1\}$ is contained in Ω. A subset Ω is a **star set** if it contains at least one star point.

Exercise: Show that $\Omega \subset \mathbb{R}^N$ is convex iff every point of Ω is a star point.

Exercise: Let Ω_1, Ω_2 be star sets.
a) Show that $\Omega_1 \cap \Omega_2$ need not be a star set.
b) If your example for part a) was one in which $\Omega_1 \cap \Omega_2 = \varnothing$, congratulations on your rigor and sense of economy. Now find another example in which the intersection is nonempty.
c) Let $\{\Omega_i\}_{i \in I}$ be a family of subsets of $\mathbb{R}^N$. Suppose that $P \in \mathbb{R}^N$ is a star point of Ω_i for all i. Show that P is a star point of $\bigcap_{i \in I} \Omega_i$.

A **central ray** in $\mathbb{R}^N$ is a subset of the form $\mathbb{R}^{\geq 0} \cdot P$ for some $P \neq 0$.

A **star body** is a subset $\Omega \subset \mathbb{R}^N$ satisfying all of the following:
(SB1) 0 is a star point of Ω.
(SB2) 0 lies in the interior of Ω.
(SB3) Every central ray intersects $\partial\Omega$ in at most one point.

A **convex body** is a convex set which is a star body.

Now in fact a convex set is a convex body iff it has zero as an interior point, i.e., this, together with convexity, implies (SB1) (obviously) and (SB3) (not so obviously). In order to get a quick, tidy proof of the latter implication, we will treat it as part of our discussion on "distance functions", coming up next.

4.4. **Distance Functions.**

Consider the following properties of a function $f : \mathbb{R}^N \to \mathbb{R}^{\geq 0}$:
(DF0) f is continuous.
(DF1) $f(x) = 0 \iff x = 0$.
(DF1′) $f(0) = 0$.
(DF2) For all $\lambda \in \mathbb{R}^+$ and all $x \in \mathbb{R}^N$, $f(\lambda x) = \lambda f(x)$.
(DF3) For all $x, y \in \mathbb{R}^N$, $f(x+y) \leq f(x) + f(y)$.
(DF4) For all $x \in \mathbb{R}^N$, $f(-x) = f(x)$.

A function satisfying (DF0), (DF1) and (DF2) is a **distance function**. A function satisfying (DF0), (DF1), (DF2) and (DF3) is a **convex distance function**. A function satisfying (DF4) is **even**. Finally, if in any of the above we replace (DF1)

with the weaker (DF1′) we speak of **pseudo-distance functions**.

Example (L^p norms on $\mathbb{R}^N$): ...

Theorem 4.6. *a) Let $f : \mathbb{R}^N \to \mathbb{R}$ be a pseudo-distance function, and put*

$$\Omega_f = f^{-1}([0,1)).$$

Then Ω_f is an open star body with boundary $f^{-1}(1)$ and exterior $f^{-1}((1,\infty))$.
b) Let Ω be a star body. We define $f_\Omega : \mathbb{R}^N \to [0,\infty)$ as follows: $f_\Omega(0) = 0$; for all $x \in \partial\Omega$, $f_\Omega(x) = 1$; for all $\lambda \in (0,\infty)$ and $x \in \mathbb{R}^N$, $f_\Omega(\lambda x) = \lambda f(x)$. Then f_Ω is a pseudo-distance function.
c) We have $f_{\Omega_f} = f$ and $\Omega_{f_\Omega} = \Omega^\circ$. In particular there is a bijective correspondence between pseudo-distance functions and open star bodies.
d) A pseudo-distance function is a distance function iff its corresponding star body is bounded.
e) If Ω is a convex set with $0 \in \Omega^\circ$, then defining f_Ω as in part b) we get a convex pseudo-distance function. Conversely, the star body corresponding to any convex pseudo-distance function is convex.
f) A pseudo-distance function is even iff its corresponding star body is centrally symmetric.

Proof. Mr. Brian A. Bonsignore gave a lecture on this material, using [C, pp. 103-111] as a source. This writeup closely follows his lecture notes.
a) Ω_f is open: by (DF0), f is continuous; and $\Omega_f = f^{-1}([0,1)) = f^{-1}((-1,1))$.
0 is a star point of Ω_f: by (DF1′), $f(0) = 0$, so $0 \in \Omega_f$. If $x \in \Omega_f$, then $f(x) < 1$, and so $f(\lambda x) = \lambda f(x) < 1$ for all $\lambda \in [0,1]$. Thus $\lambda x \in \Omega_f$ for all $\lambda \in [0,1]$.
$\partial\Omega_f = f^{-1}(1)$: if $f(x) = 1$, then for $\lambda \in [0,\infty)$, $f(\lambda x) = \lambda f(x) = \lambda$. Thus $\lambda x \in \Omega_f$ when $\lambda < 1$ and $\lambda x \notin \Omega_f$ when $\lambda > 1$. So every neighborhood of x meets Ω_f and $\mathbb{R}^N \setminus \Omega_f$, so $x \in \partial\Omega_f$. Conversely, if $f(x) < 1$, let δ be such that $f(x) < \delta < 1$. Then $x \in f^{(-1)}((-1,\delta)) \subset \Omega_f$, so $x \in \Omega_f^\circ$. The $f(x) > 1$ case is similar.
That the exterior of Ω_f is $f^{-1}((1,\infty))$ follows immediately.
b) (DF1′): Since $\lambda 0 = 0 \in \Omega$ for all $\lambda \in [0,\infty)$, $f_\Lambda(0) = 0$.
(DF2): Let $x \in \mathbb{R}^N$. If $\lambda x \in \Omega$ for all $\lambda \in [0,\infty)$, then $f_\Omega(\lambda x) = 0 = \lambda f(x)$ for all $\lambda \geq 0$. On the other hand, if $\lambda x \in \partial\Omega$ for some $\lambda > 0$, then $f_\Omega(x) = \frac{1}{\lambda}$ and $f_\Omega(\lambda x) = 1$, so

$$f_\Omega(\lambda x) = 1 = \lambda \cdot \frac{1}{\lambda} = \lambda f(x).$$

(DF0): First we show that f_Ω is continuous at 00. Fix $\epsilon > 0$. Since Ω is open, there is $\delta > 0$ such that $B(0,\delta) \subset \Omega$. If $|x| \leq \delta\epsilon$, then $\frac{\delta}{|x|}x \in \overline{B(0,\delta)} \subset \overline{\Omega}$, and thus

$$f_\Omega(x) \leq \left(\frac{\delta}{|x|}\right)^{-1} = \frac{|x|}{\delta} \leq \frac{\delta\epsilon}{\delta} = \epsilon.$$

Next let $x \in \mathbb{R}^N \setminus \{0\}$ and again fix $\epsilon > 0$. Since $\frac{1}{f_\Omega(x)+\epsilon} < \frac{1}{f_\Omega(x)}$,

$$x' = \frac{1}{f_\Omega(x)+\epsilon}x \in \Omega.$$

Since Ω is open, there is $\eta > 0$ such that $B(x',\eta) \subset \Omega$). Let $B = (f_\Omega(x)+\epsilon)B(x',\eta)$, an open neighborhood of x. If $y \in B$, then $y = (f_\Omega(x)+\epsilon)y'$ for some $y' \in B(x',\eta) \subset \Omega$, and thus

$$f_\Omega(y) = f_\Omega((f_\Omega(x)-\epsilon)y') = (f_\Omega+\epsilon)f_\Omega(y') < f_\Omega(x)+\epsilon. \tag{3}$$

If $f_\Omega(x) \leq \epsilon$, then $f_\Omega(y) \geq 0 \geq f_\Omega(x) - \epsilon$ and thus $|f_\Omega(y) - f_\Omega(x)| < \epsilon$. If $f_\Omega(x) > \epsilon$, then consider

$$x'' = \frac{1}{f_\Omega(x) - \epsilon} x \in \mathbb{R}^N \setminus \overline{\Omega}.$$

Since $\mathbb{R}^N \setminus \overline{\Omega}$ is open, there is $\eta' > 0$ such that $B(x'', \eta') \subset \mathbb{R}^N \setminus \overline{\Omega}$. Let $B'' = (f_\Omega(x) - \epsilon)B(x'', \eta')$, an open neighborhood of x. Then for $y'' \in B(x'', \eta')$ and $y = (f_\Omega(x) - \epsilon)y'' \in B''$, we have

$$f_\Omega(y) = (f_\Omega(x) - \epsilon) f_\Omega(y'') > f_\Omega(x) - \epsilon. \tag{4}$$

Thus, for $y \in B' \cap B''$, $f_\Omega(x) - \epsilon < f_\Omega(y) < f_\Omega(x) + \epsilon$ by (3) and (4).
c) First: if $f : \mathbb{R}^N \to [0, \infty)$ is a pseudo-distance function, we must show that $f_{\Omega_f} = f$. We have

$$f_{\Omega_f}(0) = 0 = f(0).$$

Suppose $x \in \mathbb{R}^N$ is such that $f(x) = 0$. Then $\lambda x \in \Omega_f$ for all $\lambda \in [0, \infty)$, so by definition of f_{Ω_f}, we have $f_{\Omega_f}(x) = 0 = f(x)$. Otherwise, $f(x) = c > 0$, and then

$$f(\frac{x}{c}) = \frac{1}{c} f(x) = 1,$$

so $\frac{x}{c} \in f^{-1}(1) = \partial\Omega_f$. Since c^{-1} is the *unique* number such that $c^{-1}x \in \partial\Omega_f$, we have $f_{\Omega_f}(x) = (c^{-1})^{-1} = c = f(x)$. Thus $f_{\Omega_f} = f$.
Next: if Ω is an open star body, we must show $\Omega = \Omega_{f_\Omega}$. By definition, $\Omega_{f_\Omega} = f_\Omega^{-1}([0,1])$. If $\lambda x \in \Omega$ for all $\lambda \in [0, \infty)$ then $f_\Omega(x) = 0$ and $x \in \Omega_{f_\Omega}$. If $x \in \Omega$ and $cx \in \partial\Omega$ for some $c > 0$, then $c > 1$ and thus $f_\Omega(x) = \frac{1}{c} < 1$, so $x \in \Omega_{f_\Omega}$. This shows $\Omega \subset \Omega_{f_\Omega}$.

Conversely, suppose $x \in \Omega_{f_\Omega}$. Then $f_\Omega(x) = 0$ or $0 < f_\Omega(x) < 1$. If $f_\Omega(x) = 0$, then $x \in \Omega$. If $0 < f_\Omega(x) < 1$, then $\frac{1}{f_\Omega(x)}$ is greater than 1 and is the unique number λ such that $\frac{1}{\lambda} x \in \partial\Omega$. Thus $x = f_\Omega(x) \frac{x}{f_\Omega(x)} \in \Omega$. This shows that $\Omega_{f_\Omega} \subset \Omega$ and thus overall that $\Omega = \Omega_{f_\Omega}$.
d) ...
e) ...
f) ... □

4.5. **Jordan measurability.**

Theorem 4.7. *Every bounded convex set is Jordan measurable.*

Proof. ... □

Exercise: a) Exhibit a bounded star body which is not Jordan measurable.
b) Exhibit a bounded star set which is not Lebesgue measurable.

5. Minkowski's Convex Body Theorem

5.1. **Statement of Minkowski's First Theorem.**

Theorem 5.1. *Let $\Omega \subset \mathbb{R}^N$ be a convex body with $\operatorname{Vol}\Omega > 2^N$. Then $\Omega \cap \mathbb{Z}^N \supsetneq \{0\}$.*

Corollary 5.2. *Let $\Omega \subset \mathbb{R}^N$ be a compact convex body with $\operatorname{Vol}\Omega = 2^N$. Then $\#(\Omega \cap \mathbb{Z}^N) > 1$.*

Proof. Let $\epsilon > 0$, and put $K_\epsilon = (1+\epsilon)\Omega$. Then for all $\epsilon > 0$, Theorem 5.1a) applies to K_ϵ, which therefore admits a nonzero lattice point. On the other hand, since K_ϵ is bounded, its intersection with $\mathbb{Z}^N$ is finite; moreover, as ϵ decreases the set $K_\epsilon \cap \mathbb{Z}^N$ either stays the same or decreases. It follows that there exists $0 \neq P \in \mathbb{Z}^N$ such that $P \in K_\epsilon$ for all $\epsilon > 0$, so P lies in $\bigcap_{\epsilon>0} K_\epsilon = \Omega$. □

Exercise: Does Corollary 5.2 remain valid when "compact" is weakened to "closed"?

5.2. **Mordell's Proof of Minkowski's First Theorem.**

Our first proof of Theorem 5.1 follows L.J. Mordell [Mo35].

Step 0: Via the rescaling $\Omega \mapsto \frac{1}{2}\Omega$, an equivalent statement is:

If a convex body $\Omega \subset \mathbb{R}^N$ has volume greater than one, it contains a nonzero point P such that $2P \in \mathbb{Z}^N$.

So let $\Omega \subset \mathbb{R}^N$ be a convex body with $\operatorname{Vol}\Omega > 1$.

Step 1: If $P, Q \in \Omega$, then by central symmetry $-Q \in \Omega$, and then by convexsity $\frac{1}{2}P + \frac{1}{2}(-Q) = \frac{1}{2}P - \frac{1}{2}Q \in \Omega$.

Step 2: For $r \in \mathbb{Z}^+$, put

$$L(r) = \#(r\Omega \cap \mathbb{Z}^N) = \#(\Omega \cap \frac{1}{r}\mathbb{Z}^N).$$

By Theorem 4.7, Ω is Jordan measurable, so Theorem 2.3 applies to show that

$$\lim_{r\to\infty} \frac{L(r)}{r^N} = \operatorname{Vol}\Omega.$$

Therefore, for sufficiently large r we must have $L(r) > r^N = \#(\mathbb{Z}/r\mathbb{Z})^N$. By the Pigeonhole Principle there exist distinct $P = (x_1, \ldots, x_n),\ Q = (y_1, \ldots, y_n) \in \mathbb{Z}^N$ such that $x_i \equiv y_i \pmod r$ for all $1 \leq i \leq n$ and $\frac{1}{r}P, \frac{1}{r}Q \in \Omega$. By Step 1,

$$R = \frac{1}{2}\left(\frac{1}{r}P\right) - \frac{1}{2}\left(\frac{1}{r}Q\right) = \frac{1}{2}\left(\frac{x_1 - y_1}{r}, \ldots, \frac{x_n - y_n}{r}\right) \in \Omega \cap \frac{1}{2}(\mathbb{Z}^N \setminus \{0\}).$$

5.3. **Statement of Blichfeldt's Lemma.**

Let us call a bounded subset $\Omega \subset \mathbb{R}^N$ **packable** if its translates by lattice points $x \in \mathbb{Z}^N$ are pairwise disjoint: i.e., for all $x \neq y \in \mathbb{Z}^N$, $(x + \Omega) \cap (y + \Omega) = \varnothing$.

Exercise: Show Ω is *not* packable iff there exist $x, y \in \Omega$ such that $x - y \in \mathbb{Z}^N \setminus \{0\}$.

Theorem 5.3. *(Blichfeldt) If $\Omega \subset \mathbb{R}^N$ is measurable and packable, then* $\operatorname{Vol}\Omega \leq 1$.

Note that in Theorem 5.3 we said "measurable", not "Jordan measurable". In fact – unlike most of the other results we have seen so far – the result holds for *Lebesgue measurable* sets. We will actually give two proofs, one using properties of Lebesgue measure, and another more elementary proof assuming that Ω is Jordan measurable and using properties of the Riemann integral.

5.4. **Blichfeldt's Lemma Implies Minkowski's First Theorem.**

But first let us demonstrate the following.

Proposition 5.4. *Blichfeldt's Lemma implies Minkowski's First Theorem.*

Proof. As above, we may assume $\operatorname{Vol}\Omega > 1$ and show $\Omega \cap (\frac{1}{2}\mathbb{Z} \setminus \{0\}) \neq \varnothing$. Applying Theorem 5.3 to Ω, there exist distinct $x, y \in \Omega$ such that $P = x - y \in \mathbb{Z}^N$. Also as above, $y \in \Omega \implies -y \in \Omega$ and thus $\frac{1}{2}x + \frac{1}{2}(-y) = \frac{1}{2}P \in \Omega \cap (\frac{1}{2}\mathbb{Z}^N \setminus \{0\})$. □

5.5. **First Proof of Blichfeldt's Lemma: Riemann Integration.**

We will prove the contrapositive: supppose Ω is packable, i.e., that the translates $\{x + \Omega \mid x \in \mathbb{Z}^N\}$ are pairwise disjoint. Let d be a positive real number such that every point of Ω lies at a distance at most d from the origin (the boundedness of Ω is equivalent to $d < \infty$).

Let $\overline{B}_r(0)$ be the closed ball of radius r centered at the origin. It has volume $c(N)r^N$ where $c(N)$ depends only on N. By Theorem 2.3 we know that the number of lattice points inside $\overline{B}_r(0)$ is asymptotic to $c(N)r^N$. Therefore the number of lattice points inside $\overline{B}_{r-d}(0)$ is asymptotic, as $r \to \infty$, to $c(N)(r-d)^N \sim c(N)r^N$. Therefore for any fixed $\epsilon > 0$, there exists R such that $r \geq R$ implies that the number of lattice points inside $\overline{B}_{r-d}(0)$ is at least $(1-\epsilon)c(N)r^N$.

Now note that if $x \in \mathbb{Z}^N$ is such that $||x|| \leq r - d$, then the triangle inequality gives $x + \Omega \subset \overline{B}_0(r)$. Then, if Ω is packable, then we have at least $(1-\epsilon)c(N)r^N$ pairwise disjoint translates of Ω contained inside $\overline{B}_0(r)$. Therefore we have

$$c(N)r^N = \operatorname{Vol}(\overline{B}_r(0)) \geq \operatorname{Vol}(P(\Omega) \cap \overline{B}_r(0)) \geq (1-\epsilon)c(N)r^N \operatorname{Vol}(\Omega),$$

and therefore

$$\operatorname{Vol}(\Omega) \leq \frac{1}{1-\epsilon}.$$

Since this holds for all $\epsilon > 0$, we conclude $\operatorname{Vol}(\Omega) \leq 1$.

5.6. **Second Proof of Blichfeldt's Lemma: Lebesgue Integration.**

Let m denote Lebesgue measure on $\mathbb{R}^N$. Suppose $\Omega \subset \mathbb{R}^N$ is Lesbesgue measurable and packable: that is the sets $\{\Omega - x\}_{x \in \mathbb{Z}^N}$ are pairwise disjoint. For each $x = (x_1, \ldots, x_n) \in \mathbb{Z}^N$, put

$$[x, x+1) = [x_1, x_1 + 1) \times \ldots \times [x_n, x_n + 1)$$

and

$$\Omega_x = \Omega \cap [x, x+1),$$

so

$$\Omega = \coprod_{x \in \mathbb{Z}^N} \Omega_x$$

and thus

$$m(\Omega) = \sum_{x \in \mathbb{Z}^N} m(\Omega_x).$$

Since Ω is packable, the family $\{\Omega_x - x\}$ is pairwise disjoint, so

$$m(\coprod_{x \in \mathbb{Z}^N} \Omega_x - x) = \sum_{x \in \mathbb{Z}^N} m(\Omega_x - x) \stackrel{*}{=} \sum_{x \in \mathbb{Z}^N} m(\Omega_x) = m(\Omega);$$

in the starred equality we have used the translation invariance of m. On the other hand, each $\Omega_x - x$ is contained in $[0,1)^N$, so

$$m(\Omega) = m(\coprod_{x \in \mathbb{Z}^N} \Omega_x - x) \leq m([0,1)^N) = 1.$$

5.7. **A Strengthened Minkowski's First Theorem.**

Let $\Lambda \subset \mathbb{R}^N$ be any full lattice. It is a simple matter to formulate a version of Minkowski's First Theorem with the standard integer lattice $\mathbb{Z}^N$ replaced by Λ.

Consider a linear automorphism $M : \mathbb{R}^N \to \mathbb{R}^N$, which we may identify with its defining matrix $M \in \mathrm{GL}_N(\mathbb{R})$ (i.e., $M = (m_{ij})$ is an $N \times N$ real matrix with nonzero determinant).

Lemma 5.5. *Let Ω be a subset of $\mathbb{R}^N$ and $M : \mathbb{R}^N \to \mathbb{R}^N$ be an invertible linear map. Consider the image*

$$M(\Omega) = \{M(x_1, \ldots, x_n)^t \mid (x_1, \ldots, x_n) \in \Omega\}.$$

a) Ω is nonempty $\iff$ $M(\Omega)$ is nonempty.
b) Ω is bounded $\iff$ $M(\Omega)$ is bounded.
c) Ω is convex $\iff$ $M(\Omega)$ is convex.
d) Ω is centrally symmetric $\iff$ $M(\Omega)$ is centrally symmetric.
e) Ω is Jordan measurable $\iff$ $M(\Omega)$ is Jordan measurable, and if so,

$$\mathrm{Vol}(M(\Omega)) = |\det(M)| \mathrm{Vol}(\Omega).$$

Exercise: Prove Lemma 5.5.

Corollary 5.6. *If $\Omega \subset \mathbb{R}^N$ is a convex body and $M : \mathbb{R}^N \to \mathbb{R}^N$ is an invertible linear map, then $M(\Omega)$ is a convex body, and $\mathrm{Vol}(M(\Omega)) = |\det(M)| \mathrm{Vol}(\Omega)$.*

Recall that the lattice points inside $r\Omega$ are precisely the $\frac{1}{r}$-lattice points inside Ω. This generalizes to arbitrary transformations as follows: for $M \in \mathrm{GL}_N(\mathbb{R})$, put

$$\Lambda := M\mathbb{Z}^N = \{M(x_1, \ldots, x_N)^t \mid (x_1, \ldots, x_N) \in \mathbb{Z}^N.$$

The map $\Lambda : \mathbb{Z}^N \to M\mathbb{Z}^N$ is an isomorphism of groups, so $M\mathbb{Z}^N$ is, abstractly, simply another copy of $\mathbb{Z}^N$. However, it is embedded inside $\mathbb{R}^N$ differently. A nice geometric way to look at it is that $\mathbb{Z}^N$ is the vertex set of a tiling of $\mathbb{R}^N$ by unit (hyper)cubes, whereas Λ is the vertex set of a tiling of $\mathbb{R}^N$ by (hyper)parallelopipeds. A single parallelopiped is called a **fundamental domain** for Λ, and the volume of a fundamental domain is given by $|\det(M)|$. We sometimes refer to the volume of the fundamental domain as the **covolume** of Λ and write

$$\mathrm{Covol}(\Lambda) = |\det(M)|.$$

Now the fundamental fact – a sort of "figure-ground" observation – is the following:

Proposition 5.7. *Let $\Omega \subset \mathbb{R}^N$ and let $M : \mathbb{R}^N \to \mathbb{R}^N$ be an invertible linear map. Then M induces a bijection between $M^{-1}(\mathbb{Z}^N) \cap \Omega$ and $\mathbb{Z}^N \cap M(\Omega)$.*

Exercise: Prove Proposition 5.7.

Applying this (with M^{-1} in place of M) gives the following: if we have a lattice $\Lambda = M\mathbb{Z}^N$, and a convex body Ω, the number of points of $\Lambda \cap \Omega$ is the same as the number of points of $\mathbb{Z}^N \cap M^{-1}(\Omega)$. Since

$$\mathrm{Vol}(M^{-1}(\Omega)) = |\det(M^{-1})| \mathrm{Vol}(\Omega) = \frac{\mathrm{Vol}(\Omega)}{\det(M)} = \frac{\mathrm{Vol}(\Omega)}{\mathrm{Covol}(\Lambda_M)},$$

we immediately deduce a more general version of Minkowski's theorem.

Theorem 5.8. *(Minkowski's First Theorem, Mark II) Let $\Omega \subset \mathbb{R}^N$ be a convex body. Let $M : \mathbb{R}^N \to \mathbb{R}^N$ be an invertible linear map, and put $\Lambda_M = M(\mathbb{Z}^N)$. Suppose that*

$$\mathrm{Vol}(\Omega) > 2^N \mathrm{Covol}(\Lambda_M) = 2^N |\det(M)|.$$

Then there exists $x \in \Omega \cap (\Lambda_M \setminus (0, \ldots, 0))$.

5.8. **Some Refinements.**

Proposition 5.9. *(Measure Theoretic Pigeonhole Principle) Let $(X, \mathcal{A}, \mu)$ be a measure space, $\{S_i\}_{i \in I}$ be a countable family of measurable subsets of X, and $m \in \mathbb{N}$. If*

$$\sum_{i \in I} \mu(S_i) > m\mu(\bigcup_{i \in I} S_i), \tag{5}$$

then there is $x \in X$ with $\#\{i \in I \mid x \in S_i\} > m$.

Proof. By replacing X with $\bigcup_{i \in I} S_i$ we may assume that $\bigcup_{i \in I} S_i = X$. Further, it is no loss of generality to assume that $\mu(X) > 0$ and that no $x \in X$ lies in infinitely many of the sets S_i: indeed, in the former case the hypothesis does not hold and in the latter case the conclusion holds.

For a subset $S \subset X$, denote by 1_S the associated characteristic function: $1_S(x) = 1$ if $x \in S$, and otherwise $1_S(x) = 0$. Put

$$f = \sum_{i \in I} 1_{S_i}.$$

For any $x \in X$, $f(x) = \#\{i \in I \mid x \in S_i\}$, so $f : X \to \mathbb{R}$ is a measurable function. The condition (5) can be reexpressed as

$$\int_X f d\mu > m \int_X d\mu,$$

so we must have $\#\{i \in I \mid x \in S_i\} = f(x) > m$ for at least one $x \in X$. □

Theorem 5.10. *(First Generalized Blichfeldt Lemma) Let $\Omega \subset \mathbb{R}^N$ be a measurable subset, let Λ be a full lattice in $\mathbb{R}^N$, and let $m \in \mathbb{Z}^+$. Suppose $\mathrm{Vol}\,\Omega > m\,\mathrm{Covol}(\Lambda)$. Then there exist distinct $w_1, \ldots, w_{m+1} \in \Omega$ such that for all $1 \leq i, j \leq m+1$, $w_i - w_j \in \Lambda$.*

Proof. For $x = (x_1, \ldots, x_n) \in \Lambda$, put

$$[x, x+1) = [x_1, x_1 + 1) \times \ldots \times [x_n, x_n + 1)$$

and

$$\Omega_x = \Omega \cap [x, x+1).$$

Then $\Omega = \coprod_{x \in \Lambda} \Omega_x$, so

$$\sum_{x \in \Lambda} m(\Omega_x - x) = \sum_{x \in \Lambda} m(\Omega_x) = m(\Omega) > m\,\mathrm{Covol}(\Lambda). \tag{6}$$

Then $\mathcal{F} = [0, 0+1)$ is a fundamental domain for Λ, so $m(\mathcal{F}) = \operatorname{Covol}\Lambda$. For all $x \in \Lambda$, $\Omega_x - x \subset \mathcal{F}$, so taking $X = \mathcal{F}$, $I = \Lambda$, $S_i = \Omega_i - i$, by (6) the hypotheses of the Measure Theoretic Pigeonhole Principle are satisfied and thus there is $v \in \mathcal{F}$ and $x_1, \ldots, x_{m+1} \in \Lambda$ such that

$$v \in \bigcap_{i=1}^{m+1} \Omega_{x_i} - x_i.$$

Thus for $1 \leq i \leq m+1$ there is $w_i \in \Omega_{x_i}$ – so $w_1, \ldots, w_{m+1}$ are distinct – such that

$$\forall 1 \leq i \leq m+1, w_i - x_i = v.$$

It follows that for all $1 \leq i, j \leq m+1$, $w_i - w_j = (x_i + v) - (x_j + v) = x_i - x_j \in \Lambda$. □

Exercise: Let $\Lambda \subset \mathbb{R}^N$ be a subgroup, $\Omega \subset \mathbb{R}^N$ a subset and $m \in \mathbb{N}$.
a) Show that the following are equivalent:
(i) There is $w \in \Omega$ such that $\#((\Omega - w) \cap \Lambda) \geq m+1$.
(ii) There is $v \in \mathbb{R}^N$ such that $\#((\Omega - v) \cap \Lambda) \geq m+1$.
(iii) There are $w_1, \ldots, w_m \in \Omega$ such that for all $1 \leq i, j \leq m$, $w_i - w_j \in \Lambda$.
(iv) There are $x_1, \ldots, x_{m+1} \in \Lambda$ such that $\bigcap_{i=1}^{m+1} \Omega + x_i \neq \varnothing$.
When these equivalent conditions hold, we say Ω is **m-packable for** Λ.
b) Show that Ω is 0-packable for Λ iff $\Omega \neq \varnothing$.
c) Show that if $\Lambda = \mathbb{Z}^N$, then Ω is 1-packable for Λ iff Ω is packable.

Exercise: Let $\Lambda \subset \mathbb{R}^N$ be a subgroup, $\Omega \subset \mathbb{R}^N$, and $m \in \mathbb{Z}^+$. We say Ω is **essentially m-packable for** Λ if for any distinct $x_1, \ldots, x_{m+1} \in \Lambda$, $m(\bigcap_{i=1}^{m+1} \Omega + x_i) = 0$.
a) Show that if a subset $\Omega \subset \mathbb{R}^N$ is m-packable for Λ then it is essentially m-packable for Λ. Give an example to show that the converse does not hold.
b) Observe that an equivalent reformulation of Theorem 5.10 is: if a measurable subset $\Omega \subset \mathbb{R}^N$ is m-packable for a lattice $\Lambda \subset \mathbb{R}^N$, then $m(\Omega) \geq m \operatorname{Covol}\Lambda$.
c) Prove the following result – which is, in view of part b), a mild strengthening of Theorem 5.10 – if $\Omega \subset \mathbb{R}^N$ is essentially m-packable for a lattice $\Lambda \subset \mathbb{R}^N$, then $m(\Omega) \geq m \operatorname{Covol}\Lambda$.

Exercise: a) Show that in under the hypotheses of the Measure Theoretic Pigeonhole Princple one can extract a stronger conclusion: there is a subset $J \subset I$ with $\#J = m+1$ such that $\mu(\bigcap_{i \in J} S_i) > 0$.
b) Observe that an equivalent reformulation of Theorem 5.10 is: suppose $\Omega \subset \mathbb{R}^N$ is Lebesgue measurable, $\Lambda \subset \mathbb{R}^N$ is a lattice, and $m \in \mathbb{Z}^+$ are such that $m(\Omega) > m \operatorname{Covol}\Lambda$. Then there is $w \in \Omega$ such that $\Omega - w$ contains at least $m+1$ points of Λ.
c) Prove the following result – which is, in view of part b), a mild strengthening of Theorem 5.10 – suppose $\Omega \subset \mathbb{R}^N$ is Lebesgue measurable, $\Lambda \subset \mathbb{R}^N$ is a lattice, and $m \in \mathbb{Z}^+$ are such that $m(\Omega) > m \operatorname{Covol}\Lambda$. Show that $\{w \in \Omega \mid \#((\Omega - w) \cap \Lambda) \geq m+1\}$ has positive Lebesgue measure.

Theorem 5.11. *(van der Corput) Let $\Omega \subset \mathbb{R}^N$ be a nonempty centrally symmetric convex subset.*
a) Then $\#(\Omega \cap \mathbb{Z}^N) \geq 2(\lceil \frac{\operatorname{Vol}(\Omega)}{2^N} \rceil - 1) + 1$.
b) If Ω is closed and bounded, then $\#(\Omega \cap \mathbb{Z}^N) \geq 2(\lfloor \frac{\operatorname{Vol}(\Omega)}{2^N} \rfloor) + 1$.

Exercise: Prove Theorems 5.10 and 5.11. (Hint/challenge: *all* of the proofs we have given so far should generalize to this context.)

5.9. **Pick's Theorem via Minkowski's Theorem.**

A convex polygon $P \subset \mathbb{R}^2$ is a **lattice polygon** if all of its vertices are points of the standard integral lattice $\mathbb{Z}^2$. The following 1899 result of G.A. Pick is a true classic: of all the theorems that I should have been taught but wasn't as a student (high school, college,...), perhaps this one is at the very top of the list.

Theorem 5.12. *(Pick's Theorem) Let $P \subset \mathbb{R}^2$ be a lattice polygon. Let I be the number of lattice points on the interior P° of P and let B be the number of lattice points on the boundary ∂P of P. Then the area of P is*

$$\mathrm{Area}(P) = I + \frac{B}{2} - 1.$$

There are many proofs of Pick's Theorem, including some which are accessible to high school students. Here we follow a recent paper of Murty and Thain [MuTh07] which deduces Pick's Theorem from Minkowski's Convex Body Theorem.

Proof. Step 0: Define a **elementary triangle** to be a lattice triangle with no interior lattice points. Pick's Theorem applied to an elementary triangle is precisely the assertion that the area of any elementary triangle is $\frac{1}{2}$. Our strategy is as follows: first we will prove Pick's Theorem for elementary triangles and then we will deduce the general case of Pick's Theorem from this by an induction / dissection argument.
Step 1: We claim that the area of any elementary triangle is *at least* $\frac{1}{2}$. Indeed, the area of triangle ABC is equal to half the area of the parallelogram with sides AB, BC and thus half the magnitude of the cross product $(A-B) \times (C-B)$. But the cross product of $A-B$ and $C-B$ is a nonzero lattice vector, so its length is at least 1, and thus the area of ABC is at least $\frac{1}{2}$.
Step 2: We claim the area of any elementary triangle T_1 is *at most* $\frac{1}{2}$. Let the vertices of the triangle T_1 be ABC. Each of the following linear transformations sends lattice points to lattice points: 180° rotation about a lattice point, and translation by a lattice point. Thus the image of any elementary triangle under any composition of such transformations is again an elementary triangle. Consider the triangle T_2 obtained by reflecting T_1 through the line AB. Then T_2 can also be obtained by rotating 180° around vertex B and translating by the lattice point $A-C$, so it is an elementary triangle. Similarly, we may attach elementary triangles T_2, T_3, T_4 by reflecting across the other two sides of the triangle. This resulting union of four elementary triangles is again an elementary triangle, say P_1. Finally, consider P_2, the elementary triangle obtained by rotating P_1 180° about P_1, and consider $\Omega = P_1 \cup P_2$. Then Ω° is a convex body which is symmetric around the lattice point B has no nonzero lattice points, and has area $8\,\mathrm{Area}(T_1)$. If $\mathrm{Area}(T_1) > \frac{1}{2}$ then $\mathrm{Area}(\Omega^\circ) > 4$, contradicting Minkowski's Convex Body Theorem.
Step 3: We may *triangulate* P, i.e., by drawing straight lines connecting some of the vertices, we express P as a finite union of lattice triangles $T_1 \cup T_n$, such that for all $i \neq j$, T_i and T_j are either disjoint or share precisely one common side. And we may go further: for each interior lattice point P of a triangle $T_i = ABC$ we may draw line segments AP, BP, CP, which subdivides T_i into three lattice triangles and decreases the total number of interior lattice points by one. Applying this

procedure finitely many times we arrive at a triangulation of P into finitely many **elementary triangles**, i.e., lattice triangles possessing no interior lattice points.

We now claim that for all $n \in \mathbb{Z}^+$, every convex lattice polygon which is a union of n elementary triangles satisfies Pick's Theorem. We prove this by induction on n: the base case – i.e., 1 elementary triangle – has been shown in Steps 1 and 2 above. Now suppose that Pick's Theorem holds for all convex lattice polygons which are a union of n elementary triangles, and let P_{n+1} be a convex lattice polygon which is a union of $n+1$ elementary triangles. Take one of the exterior elementary triangles T and write $P_{n+1} = P_n \cup T$. Let I_n and B_n be the number of interior and boundary lattice points for P_n, and let I_{n+1}, B_{n+1} be these same quantities for P_{n+1}. Then we have $I_{n+1} = I_n$, $B_{n+1} = B_n + 1$ and $\text{Area}(P_{n+1}) = \text{Area}(P_n) + \frac{1}{2}$, so

$$\text{Area}(P_{n+1}) = \text{Area}(P_n) + \frac{1}{2} = I_n + \frac{B_n}{2} - 1 + \frac{1}{2} = I_{n+1} + \frac{B_{n+1}}{2} - 1. \qquad \square$$

Exercise: Show there are infinitely many noncongruent elementary triangles.

Exercise: Draw nice pictures for the argument in Step 2 of the proof of Theorem 5.12. (In [MuTh07] they draw only a single diagram, and their basic triangle ABC appears to be an isosceles right triangle: the result is trivial in this case!)

Exercise: a) Show that the hypothesis of convexity of the lattice polygon P can be relaxed to **simplicity**: i.e., plane regions whose boundary consists of a single simple polygonal curve.
b) An example of a polygonal region to which Pick's Theorem *does not* apply is $[0,3] \times [0,3] \setminus (1,2) \times (1,2)$, i.e., a "square torus". Formulate a more general version of Pick's Theorem which applies to such regions. (Suggestion: replace the "1" which occurs in Pick's Theorem with the **Euler characteristic** of P.)

6. Minkowski's Theorem on Successive Minima

Let $\Omega \subset \mathbb{R}^N$ be a convex body. Define the **Minkowski minimum** $\lambda_1 = \lambda_1(\Omega)$ to be the infimum over all positive real numbers λ such that $\dim_{\mathbb{R}} \langle (\lambda\Omega) \cap \mathbb{Z}^N \rangle_{\mathbb{R}} \geq 1$.

Equivalently but more simply, λ_1 is the smallest dilate of Ω which contains a nonzero lattice point. Since $\text{Vol}(\lambda\Omega) = \lambda^N \text{Vol}\,\Omega$, by Minkowski's First Theorem we certainly have a nontrivial lattice point when $\lambda > \frac{2}{(\text{Vol}\,\Omega)^{\frac{1}{n}}}$, and thus

$$\lambda_1(\Omega) \leq \frac{2}{(\text{Vol}\,\Omega)^{\frac{1}{n}}}. \tag{7}$$

Exercise: a) Show that Minkowski's First Theorem is equivalent to (7).
b) State a version of (7) with an arbitrary lattice Λ in place of $\mathbb{Z}^N$.

The point of this somewhat unlikely looking restatement of Minkowski's First Theorem is the following generalization. For $1 \leq i \leq n$, we define the **ith successive minimum** $\lambda_i = \lambda_i(\Omega)$ as

$$\lambda_i = \inf\{\lambda \in \mathbb{R} \mid \dim_{\mathbb{R}} \langle (\lambda\Omega) \cap \mathbb{Z}^N \rangle_{\mathbb{R}} \geq i.$$

That is, λ_i is the least dilate of Ω needed to attain not just a nonzero lattice point but i $\mathbb{R}$-linearly independent lattice points. Clearly we have

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n.$$

Moreover we may replace $\mathbb{Z}^N$ by a lattice Λ in $\mathbb{R}^N$ and define successive minima

$$\lambda_1(\Omega, \Lambda) \leq \lambda_2(\Omega, \Lambda) \leq \ldots \leq \lambda_n(\Omega, \Lambda).$$

Theorem 6.1. *(Minkowski's Second Theorem) Let $\Omega \subset \mathbb{R}^N$ be a convex body and $\Lambda \subset \mathbb{R}^N$ a full lattice. Then:*

$$\frac{1}{n!} \prod_{i=1}^{N} \frac{2}{\lambda_i(\Omega, \Lambda)} \leq \frac{\operatorname{Vol} \Omega}{\operatorname{Covol} \Lambda} \leq \prod_{i=1}^{N} \frac{2}{\lambda_i(\Omega, \Lambda)}. \tag{8}$$

7. The Minkowski-Hlawka Theorem

Minkowski's Convex Body Theorem asserts that a convex body in $\mathbb{R}^N$ of volume greater than 2^N must have a nonzero point in every lattice Λ of covolume one (and the number 2^N is best possible as we range over all convex bodies). It is natural to ask about the converse: suppose that a subset Ω has a nonzero point in every lattice of covolume one. Does this imply a *lower bound* on the volume of Ω?

Minkowski himself conjectured an answer. More precisely, Minkowski's seminal writings on geometry of numbers contain the statement of a theorem of this type, but no proof was ever published by Minkowski or found in the unpublished work he left behind after his death in 1909. Minkowski's "theorem" was first proved in 1943 in a celebrated work of E. Hlawka [Hl43]. Because (or in spite of?) this history it is traditional to speak of the **Minkowski-Hlawka Theorem**.

7.1. Statement of the theorem.

Theorem 7.1. *(Minkowski-Hlawka) Let $N \geq 2$, and let $\Omega \subset \mathbb{R}^N$ be bounded and Jordan measurable. Suppose that Ω has a nonzero point in every lattice $\Lambda \subset \mathbb{R}^N$ of covolume 1.*
a) Then $\operatorname{Vol} \Omega \geq 1$.
b) If Ω is a star body, then $\operatorname{Vol} \Omega \geq \zeta(N) = \sum_{k=1}^{\infty} \frac{1}{k^N}$.
c) If Ω is a symmetric star body, then $\operatorname{Vol} \Omega \geq 2\zeta(N)$.

Exercise: Show that the results of Theorem 7.1 are false for $N = 1$.

Many different proofs of Theorem 7.1 (and also improvements, variants,...) have been published over the years:[5] [Hl43], [Ma44] (with $\zeta(N)$ replaced by $\frac{1}{N}$ – i.e., a quantitatively weaker result), [Si45] (as a consequence of a **Mean Value Theorem** in the geometry of numbers), [Ma46] (with a *better* constant than $\zeta(N)$ in the case of a convex body), [DaRo47] (improvements on the preceding paper), [Ro47], [Ro51] (by a method which is related to van der Corput's extension of Minkowski's First Theorem), [Ca53a], [Ro56] (a substantial quantitative improvement for large N), [Sa68] (a version of Minkowski-Hlawka with unimodular matrices replaced by rotation matrices) and [Th96] (an adelic analogue).

[5]What follows is **not** a comprehensive list!

7.2. **Proof of Minkowski-Hlawka, Part a).**

The following argument is taken from Hardy and Wright.[6] Hardy and Wright are, in turn, following Rogers [Ro47], but (as usual) I find their exposition to be superior.

Let $N \geq 2$ and let $\Omega \subset \mathbb{R}^N$ be bounded and Jordan measurable, with $\operatorname{Vol}\Omega < 1$. We need to find a covolume 1 lattice $\Lambda \subset \mathbb{R}^N$ such that $\Lambda^\bullet \cap \Omega = \emptyset$.

Step 1: Let $C > 0$ be such that $\Omega \subset [-C, C]^N$. Let $p > C^N$ be a prime number. Let $A_1, \ldots, A_{N-1} \in \mathbb{Z}$ and put $A = (A_1, \ldots, A_{N-1})$. For each such A, we define a lattice $\Lambda_A = M_A \mathbb{Z}^N$, with

$$M_A = \begin{bmatrix} \frac{1}{p^{\frac{N-1}{N}}} & 0 & 0 & \ldots & 0 \\ \frac{A_1}{p^{\frac{N-1}{N}}} & \frac{p}{p^{\frac{N-1}{N}}} & 0 & \ldots & 0 \\ \frac{A_2}{p^{\frac{N-1}{N}}} & 0 & \frac{p}{p^{\frac{N-1}{N}}} & \ldots & 0 \\ \vdots & & & & \\ \frac{A_{N-1}}{p^{\frac{N-1}{N}}} & 0 & 0 & \ldots & \frac{p}{p^{\frac{N-1}{N}}} \end{bmatrix}.$$

Then $\det M_A = 1$, so $\operatorname{Covol}\Lambda_A = 1$. The strategy of the proof is as follows: we will assume that for all $A \in \mathbb{Z}^{N-1}$, $\Lambda_A^\bullet \cap \Omega \neq \varnothing$ and deduce $\operatorname{Vol}\Omega \geq 1$.

Step 2: So, suppose that for each $A \in \mathbb{Z}^{N-1}$ there exists a nonzero point $P_A \in \Lambda_A \cap \Omega$. Such a point is of the form $M_A(x_1, \ldots, x_N)^T$ for $(x_1, \ldots, x_N) \in (\mathbb{Z}^N)^\bullet$, so

$$P_A = (P_1, \ldots, P_N) = M_A(x_1, \ldots, x_N)^T = \frac{1}{p^{\frac{N-1}{N}}}(x_1, A_1 x_1 + p x_2, \ldots, A_{n-1} x_1 + p x_N).$$

Suppose $x_1 = 0$. Then for $1 \leq i \leq N-1$ we have

$$|P_i| = p^{\frac{1}{N}} |x_{i+1}| \leq C < p^{\frac{1}{N}},$$

so $x_2 = \ldots = x_N = 0$ and thus $P_A = 0$, contradiction. So $x_1 \neq 0$, and moreover

$$0 < |x_1| = p^{\frac{N-1}{N}} |P_1| \leq C p^{\frac{N-1}{N}} < p^{\frac{1}{N}} p^{\frac{N-1}{N}} = p.$$

Next suppose that for $A, A' \in \mathbb{Z}^{N-1}$ we have $P_A = P_{A'} = M_{A'}(x_1', \ldots, x_N')^T$. Then $x_1 = x_1'$ and thus for all $1 \leq i \leq N-1$,

$$A_i x_1 + p x_{i+1} = A_i' x_1 + p x_{i+1}'.$$

Since $\gcd(x_1, p) = 1$, it follows that for all i, $p \mid (A_i - A_i')$.
Step 3: Because of Step 2, the map $\{0, \ldots, p-1\}^{N-1} \to \Omega$ given by $A = (A_1, \ldots, A_{N-1}) \mapsto P_A$ is injective. This shows

$$p^{N-1} \leq \# \left(\frac{1}{p^{\frac{N-1}{N}}} \mathbb{Z}^N \cap \Omega \right) = \#(\mathbb{Z}^N \cap p^{\frac{N-1}{N}} \Omega) = L_\Omega(p^{\frac{N-1}{N}}),$$

[6]They present their proof with $N = 2$ and remark that it extends "at once" to N dimensions. It took me about half an hour to figure out how to make this extension: close enough, I suppose.

where $L_\Omega(r)$ is the lattice point enumerator of §1. Equivalently, for all $p > C^N$,

$$\frac{L_\Omega(p^{\frac{N-1}{N}})}{(p^{\frac{N-1}{N}})^N} \geq 1. \tag{9}$$

Since Ω is bounded and Jordan measurable, by Theorem 2.3 we have

$$\lim_{r \to \infty} \frac{L_\Omega(r)}{r^N} = \operatorname{Vol} \Omega. \tag{10}$$

Comparing (9) and (10)[7], we deduce $\operatorname{Vol} \Omega \geq 1$.

7.3. **Primitive Lattice Points.**

Let $(G, +)$ be a commutative group. A nonzero element $x \in G$ is **primitive** if it is not nontrivially divisible in G: precisely, x is primitive iff for all $y \in G$ and $n \in \mathbb{Z}^+$, $ny = x \implies n = \pm 1$.

Convention: For the applications to follow, it turns out to be convenient to regard the identity element 0 as a primitive element, even though the definition seems to exclude it. (No big deal either way, of course...)

Exercise: a) Let x be a primitive element of the commutative group G. Show that x has infinite order.
b) Show that the group $(\mathbb{Q}, +)$ has no primitive elements. (More generally, this holds for any **divisible** abelian group.)
c) Show that the primitive elements of $(\mathbb{Z}, +)$ are precisely ± 1.

On the other hand, for $N > 1$ the group $\mathbb{Z}^N$ has infinitely many primitive elements, for instance $(1, \ldots, 1, a)$ for any $a \in \mathbb{Z}$. It is easy to give an algebraic characterization of primitive elements in $\mathbb{Z}^N$.

Proposition 7.2. $(x_1, \ldots, x_N) \in (\mathbb{Z}^N)^\bullet$ *is primitive iff* $\gcd(x_1, \ldots, x_N) = 1$.

Exercise: Prove Proposition 7.2.

There is an interesting **geometry of primitive vectors** in $\mathbb{Z}^N$ which gives rise to an important generalization of our lattice point enumerator function.

Namely, suppose that $\Lambda \subset \mathbb{R}^N$ is a lattice. Then as an additive group, $\Lambda \cong (\mathbb{Z}^N, +)$ so the above characterization of primitive elements applies: let $e_1, \ldots, e_N$ be a $\mathbb{Z}$-basis for Λ, so an arbitrary $v \in \Lambda$ has a unique expression as $v = x_1 e_1 + \ldots + x_N e_N$. Really this gives an isomorphism $\Lambda \xrightarrow{\sim} \mathbb{Z}^N$, which shows that v is primitive iff $\gcd(x_1, \ldots, x_N) = 1$.

There is an alternate, more geometric take on primimtive points in lattices which is worth mentioning. Namely, let v_1, v_2 be two vectors in a lattice $\Lambda \subset \mathbb{R}^N$. We say that v_2 is **visible from** v_1 if the line segment joining v_1 to v_2 does not contain any other points of Λ. Of course v_2 is visible from v_1 iff v_1 is visible from v_2.

[7]And using the fact that there are infinitely many prime numbers!

Proposition 7.3. *Let $\Lambda \subset \mathbb{R}^N$ be a lattice, and let $v, v_1, v_2 \in \Lambda$.*
a) Then v_2 is visible from v_1 iff $v_2 - v_1$ is a primitive vector.
b) In particular, a vector $v \in \Lambda$ is primitive iff it is visible from the origin 0.

Exercise: Prove Proposition 7.3.

Proposition 7.3 suggests that the visibility relation is simply a repackaging of the notion of primitive vectors. For our purposes this will be true, but we should remark that by making one further definition one gets an array of interesting problems in discrete geometry of a kind distinct from those we are considering here. Namely, suppose S_1 and S_2 are subsets of $\mathbb{R}^N$. Then we say S_1 **is visible from** S_2 if for each $x \in S_1$ there exists $y \in S_2$ such that the line segment $\overline{xy}$ contains no elements of S_1 other than x and (possibly) y.[8] There is a rapidly increasing literature on problems involving this concept.

Let us consider the **primitive lattice point enumerator function**: for $\Omega \subset \mathbb{R}^N$ a bounded set and $r \in \mathbb{R}^{>0}$, let $\mathrm{PL}_\Omega(r)$ be the number of primitive points in the standard integer lattice $\mathbb{Z}^N$ which are contained in the dilate $r\Omega$. More explicitly, we are counting one plus the number of $(x_1, \ldots, x_N) \in \mathbb{Z}^N$ such that $\gcd(x_1, \ldots, x_N) = 1$ and $(\frac{x_1}{r}, \ldots, \frac{x_N}{R}) \in \Omega$.

Let $\zeta(s)$ be the **Riemann zeta function**, defined for $\Re s > 1$ by $\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s}$ and having a meromorphic continuation to $\mathbb{C}$ with a single simple pole at $s = 1$.

Theorem 7.4. *Let $N \geq 2$ and $\Omega \subset \mathbb{R}^N$ bounded and Jordan measurable. Then*

$$\lim_{r \to \infty} \frac{PL_\Omega(r)}{r^N} = \frac{\mathrm{Vol}\,\Omega}{\zeta(N)}.$$

Proof. **Exercise** for Lauren Huckaba and Alex Rice: see [HW6ed, §24.10]. The proof makes use of some (simple) results of §16.5 and §17.5 of *loc. cit.* concerning analytic aspects of Möbius inversion. □

Just as we did in §1 for L_Ω, it is very natural to consider the error function

$$\mathrm{PE}_\Omega(r) = |L_\Omega(r) - \left(\frac{\mathrm{Vol}\,\Omega}{\zeta(N)}\right) r^N|$$

and try to get upper bounds in terms of smoothness conditions on $\partial\Omega$. Plenty of work has indeed been done on this – though not as much for the conventional lattice point enumerator function – and it seems that this problem has a more number-theoretic flavor than the corresponding one for the conventional lattice point enumerator. In particular, much better bounds on PE_Ω can be attained under the assumption of the **Riemann Hypothesis**.

There is a lot of neat stuff here and plenty of other possible directions for generalization – what if we tried to enumerate lattice points satisfying some other "arithmetic condition" besides primitivity? – but for now we move on to complete the proof of Minkowski-Hlawka.

[8]The literature I have consulted does not make it clear whether we should allow $y \in S_1$ or not. Probably both cases should be considered...

7.4. **Proof of Minkowski-Hlawka Part b).**

Let $\Omega \subset \mathbb{R}^N$ be a bounded, Jordan measurable star body such that for every covolume one lattice $\Lambda \subset \mathbb{R}^N$, we have $\Omega \cap \Lambda^\bullet \neq \varnothing$. We wish to show that $\operatorname{Vol} \Omega \geq \zeta(N)$.

The basic setup is identical to that of part a): especially, we continue to use the lattice Λ_A for $A = (A_1, \ldots, A_{N-1}) \in \mathbb{Z}^{N-1}$. The extra idea here is as follows: the star body structure of Λ allows us to use the primitive lattice point enumerator PL_Ω rather than the conventional lattice point enumerator L_Ω, and thus the extra factor of $\zeta(N)$ comes directly from Theorem 7.4.

How does this work? It's very simple: for any lattice $\Lambda \subset \mathbb{R}^N$, let $v \in \Lambda$. If $e_1, \ldots, e_N$ is a $\mathbb{Z}$-basis of Λ, we may uniquely write $v = x_1 e_1 + \ldots + x_N e_N$ for $x_i \in \mathbb{Z}$. Put $d = \gcd(x_1, \ldots, x_N)$. Then v is primitive iff $d = 1$, but in every case we can canonically associate to v a primitive lattice point

$$v' := \frac{1}{d} P.$$

Now for all $A \in \mathbb{Z}^{N-1}$, since $P_A \in \Omega \cap \Lambda_A^\bullet$ and Ω is a star body, then also $P'_A = \frac{1}{d} P_A \in \Omega \cap \Lambda_A^\bullet$. We CLAIM that, as above, for $(A_1, \ldots, A_{N-1})$ in $\{0, \ldots, p-1\}^{N-1}$, the points P'_A are pairwise distinct. Indeed, let

$$P_A = \frac{1}{p^{N-1} N}(x_1, A_1 x_1 + p x_2, \ldots, A_{n-1} x_1 + p x_n),$$

$$P_B = \frac{1}{p^{N-1} N}(y_1, B_1 y_1 + p y_2, \ldots, B_{n-1} y_1 + p y_n),$$

and suppose that $P'_A = P'_B$. Then there exist $d_1, d_2 \in \mathbb{Z}^+$ such that $\gcd(x_1, A_1 x_1 + p x_2, \ldots, A_{n-1} x_1 + p x_n) = d_1$, $\gcd(y_1, B_1 x_1 + p y_2, \ldots, B_{n-1} y_1 + p y_n) = d_2$ and

$$d_2(x_1, A_1 x_1 + p x_2, \ldots, A_{n-1} x_1 + p x_n) = d_1(y_1, B_1 y_1 + p y_2, \ldots, B_{n-1} y_1 + p y_n). \tag{11}$$

Since $\gcd(x_1 y_1, p) = 1$ and $d_1 d_2 \mid x_1 y_1$, $\gcd(d_1 d_2, p) = 1$. Equating the first two coordinates of (11) gives

$$d_2 x_1 = d_1 y_1$$

$$d_2 A_1 x_1 + d_2 p x_2 = d_1 B_1 y_1 + d_1 p y_2 = d_2 x_1 B_1 + d_1 p y_2.$$

$$(A_1 - B_1) d_2 x_1 = p(d_2 x_2 - d_1 y_2),$$

and since $\gcd(p, d_2 x_1) = 1$, we conclude $p \mid A_1 - B_1$ and thus, since $0 < A_1, B_1 < p$, $A_1 = B_1$. In a similar way we deduce $A_2 = B_2, \ldots, A_N = B_N$, so $P_A = P_B$.

Thus for all $p > C^N$ we have

$$\frac{\mathrm{PL}_\Omega(p^{\frac{N-1}{N}})}{(p^{\frac{N-1}{N}})^N} \geq 1. \tag{12}$$

On the other hand, by Theorem 7.4 we have

$$\lim_{r \to \infty} \frac{\mathrm{PL}_\Omega(r)}{r^N} = \frac{\operatorname{Vol} \Omega}{\zeta(N)}. \tag{13}$$

Comparing (12) and (13) we deduce $\operatorname{Vol} \Omega \geq \zeta(N)$.

7.5. **Proof of Minkowski-Hlawka Part c).**

Finally, we suppose that Ω is a *symmetric* star domain. Then, for each primitive lattice point $P'_A \in \Omega$, its negative $-P'_A$ is also a primitive lattice point in Ω. If we can check that $-P'_A \neq P'_B$ for any $B \in \{0, \ldots, p-1\}^{N-1}$, then we have found twice as many primitive lattice points as we had before, and feeding this into the primitive lattice point enumerator function $\mathrm{PL}_\Omega(r)$ as in part b) above we extract an extra factor of 2 in $\mathrm{Vol}(\Omega)$. But in fact for all $B = (B_1, \ldots, B_{N-1}) \in \{0, \ldots, p-1\}^{N-1}$, the first coordinate of P'_B is positive, so of course it is not equal to $-P'_A$: done.

A more geometric / conceptual take on the above arugment is as follows: we cut our symmetric star body Ω by the hyperplane $H = \{x_1 = 0\}$. Indeed for any symmetric subset $\Omega \subset \mathbb{R}^N$ and any hyperplane H, if we put

$$\Omega^+ = \{x \in \Omega \mid H(x) > 0\}, \Omega^- = \{x \in \Omega \mid H(x) < 0\},$$

then $x \mapsto -x$ gives an isometry $\Omega^+ \to \Omega^-$: if $H(x) > 0$, then $H(-x) = -H(x) < 0$. Thus, if Ω is a bounded symmetric star body, then Ω^+ and Ω^- are star bodies with

$$\mathrm{Vol}(\Omega^+) = \mathrm{Vol}(\Omega^-) = \frac{1}{2}\mathrm{Vol}(\Omega)$$

and for any lattice Λ,

$$(\#\Lambda \cap \Omega^+) + (\#\Lambda \cap \Omega^-) + (\#\Lambda \cap \Omega \cap \{H = 0\}) = \#\Lambda \cap \Omega.$$

Thus, as Neil Lyall suggests, if we can find a hyperplane H which does not contain any nonzero lattice points in Ω, then we immediately deduce from part b) that $\mathrm{Vol}\,\Omega^+ \geq \zeta(N)$ and thus $\mathrm{Vol}\,\Omega = 2\,\mathrm{Vol}\,\Omega^+ \geq 2\zeta(N)$. This is easily done:

Exercise: Let $\Lambda \subset \mathbb{R}^N$ be a lattice, and let $\Omega \subset \mathbb{R}^N$ be a bounded set.
a) Show that for any hyperplane H_0, there exist arbitrarily close hyperplanes H (as measured using the unit normal vector, for instance) H such that $H \cap \Lambda^\bullet \cap \Omega = \varnothing$.
b) Let $c_1, \ldots, c_{N-1}$ be real numbers, not all zero. Show that if $c_N \in \mathbb{R}$ is such that there exists $v = (x_1, \ldots, x_N) \in \Lambda^\bullet$ such that $c_1x_1 + \ldots + c_{N-1}x_{N-1} + c_Nx_N = 0$, then c_N lies in the finite-dimensional $\mathbb{Q}$-vector space generated by $c_1, \ldots, c_{N-1}$. Deduce from this that for fixed $c_1, \ldots, c_{N-1}$, for all but countably many $c_N \in \mathbb{R}^N$ the hyperplane $H(x) = c_1x_1 + \ldots + c_Nx_N$ meets Λ only at 0.
c) Prove or disprove: The set of hyperplanes H in $\mathbb{R}^N$ such that $H \cap \Lambda \supsetneq \{0\}$ has measure zero in an appropriate sense (e.g. as a subset of $\mathbb{P}^N(\mathbb{R})$).

Finally, Alex Rice suggests that in fact *any* hyperplane H will work: if Ω is bounded and Jordan measurable in $\mathbb{R}^N$ then $\Omega \cap H$ is bounded and Jordan measurable in H, viewed as an $N-1$-dimensional Euclidean space and thus $\#L_{\Omega\cap H}(r) = O(r^{N-1}) = o(r^N)$. Thus the number of lattice points on the hyperplane $H = 0$ is asymptotically negligible and the argument goes through.

8. Mahler's Compactness Theorem

9. Lattice Points in Star Bodies

9.1. L^p **norms.**

For complex numbers z with positive real part, the **gamma function** is defined as

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt.$$

Proposition 9.1. *Let $n \in \mathbb{N}$. Then:*

$$\Gamma(n+1) = n! \tag{14}$$

and

$$\Gamma(n+\frac{1}{2}) = \sqrt{\pi}\left(\frac{1\cdot 3\cdots(2n-1)}{2^n}\right) = \sqrt{\pi}\left(\frac{(2n)!}{2^{2n}n!}\right). \tag{15}$$

Exercise: Prove Proposition 9.1.

For future reference we give exact values and decimal approximations for some small values of Γ at positive half/integer arguments.

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \approx 1.7724538509055160273$$

$$\Gamma(1) = 0! = 1,$$

$$\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2} \approx 0.8862269254527580137$$

$$\Gamma(2) = 1! = 1,$$

$$\Gamma\left(\frac{5}{2}\right) = \frac{3\sqrt{\pi}}{4} \approx 1.3293403881791370205$$

$$\Gamma(3) = 2! = 2,$$

$$\Gamma\left(\frac{7}{2}\right) = \frac{15\sqrt{\pi}}{8} \approx 3.3233509704478425512\,,$$

$$\Gamma(4) = 3! = 6.$$

Theorem 9.2. *Let $N \in \mathbb{Z}^+$ and $p \geq 1$. Consider the even distance function*

$$f_p : \mathbb{R}^N \to \mathbb{R}^{\geq 0}, x \mapsto \left(\sum_{i=1}^N |x_j|^p\right)^{\frac{1}{p}}.$$

Let $\mathcal{B}_{N,r} = f_p^{-1}([0,1))$ be the open unit ball in $(\mathbb{R}^N, ||\ ||_p)$. Then

$$\mathrm{Vol}(\mathcal{B}_{N,r}) = \frac{2^N\Gamma(\frac{1}{p}+1)^N}{\Gamma(\frac{N}{p}+1)}.$$

In particular, taking $r = 2$, the volume of the Euclidean unit ball in $\mathbb{R}^N$ is

$$V_N = \frac{2\pi^{N/2}}{N\Gamma(\frac{N}{2})}.$$

Proof. [S, pp. 25-26]. □

In particular we have

$$V_2 = \pi,\ V_3 = \frac{4\pi}{3},\ V_4 = \frac{\pi^2}{2}.$$

9.2. **Linear Forms.**

9.2.1. *Minkowski's Linear Forms Theorem.*

Theorem 9.3. *Let $\Lambda \subset \mathbb{R}^N$ be a lattice. Let $C = (c_{ij}) \in M_N(\mathbb{R})$ be a matrix. Consider the associated system of* ***linear forms***

$$L_i(x) = L_i(x_1, \ldots, x_n) = \sum_{j=1}^{n} c_{ij} x_j, 1 \leq i \leq n.$$

Let $\epsilon_1, \ldots, \epsilon_n$ be positive real numbers such that

$$|\det C| \operatorname{Covol} \Lambda \leq \prod_{i=1}^{N} \epsilon_i. \tag{16}$$

Then there is $x = (x_1, \ldots, x_N) \in \Lambda^\bullet$, with $|L_i(x)| \leq \epsilon_i$ for all $1 \leq i \leq N$.

Proof. Step 0: If L is a linear form, the function $x \mapsto |L(x)|$ is a symmetric, convex distance function, so that the sets $\{x \mid |L(x)| < \epsilon\}$ are symmetric convex bodies.
Step 1: Suppose that $|\det C| > 0$. By replacing L_i with $\frac{L_i}{\epsilon_i}$ one sees it is no loss of generality to assume $\epsilon_1 = \ldots = \epsilon_N = 1$. Now consider

$$\Omega = \{x \in \mathbb{R}^N \mid \forall 1 \leq i \leq N, |L_i(x)| \leq 1\}.$$

Then Ω is a compact, symmetric convex body. Indeed it is the image of the cube $[-1,1]^N$ under the linear transformation C^{-1}, so it has volume $\frac{\operatorname{Vol}([-1,1]^N}{\det C} = \frac{2^N}{\det C}$. Therefore the inequality (16) is equivalent to $\operatorname{Vol} \Omega \geq 2^N \operatorname{Covol} \Lambda$, so Minkowski's Convex Body Theorem applies to give a point $x \in \Lambda^\bullet \cap \Omega$, the desired result.
Step 2: When $\det C = 0$, the region Ω is a symmetric convex body of infinite volume, so the result holds for any positive $\epsilon_1, \ldots, \epsilon_n$. □

Remark 4. *Our earlier observation that Minkowski's Convex Body Theorem is* ***sharp*** *in the sense that the constant 2^N cannot be improved can be repeated at this point to observe that the constant in (16) is sharp: indeed, if $L(x_i) = x_i$ for all i and $\epsilon_i < 1$ for all i, then there is of course no $x \in (\mathbb{Z}^N)^\bullet$ such that $|L_(x_i)| < \epsilon_i$ for all i.*

Exercise: Show that in Theorem 9.3 the generality of arbitrary matrix C *and* also an arbitrary lattice Λ is illusory: i.e., deduce Theorem 9.3 from the special cases:
(i) $L_i(x) = x_i$ for all i and
(ii) $\Lambda = \mathbb{Z}^n$.

Exercise: Let $C = (c_{ij}) \in \mathrm{GL}_N(R)$; for $1 \leq i \leq N$ put $L_i(x) = \sum_{j=1}^{n} c_{ij} x_j$. Show: $\exists\, x = (x_1, \ldots, x_N) \in \mathbb{Z}^N \setminus \{0\}$ such that $|L_i(x)| \leq |\det C|^{\frac{1}{N}}$ for all $1 \leq i \leq N$.

9.3. **Products of Linear Forms.**

Let $C = (c_{ij}) \in \mathrm{GL}_N(\mathbb{R})$, and consider again the associated system of linear forms

$$L_i(x) = \sum_{j=1}^{N} c_{ij} x_j,$$

and put

$$D = |\det C|.$$

By Theorem 9.3, for any $\epsilon_1, \ldots, \epsilon_N > 0$ such that $D \leq \prod_{i=1}^N \epsilon_i$, there is $x \in (\mathbb{Z}^N)^\bullet$ with $|L_i(x)| \leq \epsilon_i$ for all $1 \leq i \leq N$. Now consider $f(x) = L_1(x) \cdots L_N(x)$ the product of N linear forms. Thus there is $x \in (\mathbb{Z}^N)^\bullet$ such that

$$|f(x)| = |L_1(x) \cdots L_N(x)| \leq D. \tag{17}$$

Exercise: Show that (9.4) is sharp when $N = 1$.

Note that for $N > 1$, the level sets of $|f|$ are non-convex starbodies, so Minkowski's Theorem does not apply to them directly. We can bring MCBT to bear however by finding an **inscribed convex body**. Let $g(x) = \frac{1}{N} \sum_{i=1}^N |x|$. Thus g is a rescaling of the L^1-norm on $\mathbb{R}^N$, so it is a symmetric distance function, and by Theorem 9.2, if $\Omega = g^{-1}([0,1))$ is the corresponding convex body,

$$\mathrm{Vol}(\Omega) = N^N \mathrm{Vol}(\mathcal{B}_{N,1}) = \frac{2^N N^N}{N!}.$$

Making a linear change of variables, we get

$$\mathrm{Vol}((g \circ C)^{-1}([0,1])) = \mathrm{Vol}(C^{-1} \circ g^{-1}([0,1)) = \mathrm{Vol}(C^{-1}\Omega) = \frac{\mathrm{Vol}\,\Omega}{D} = \frac{2^N N^N}{DN!}.$$

By the AGM inequality, for $x \in \mathbb{R}^N$,

$$f(x) = |L_1(x) \cdots L_N(x)| \leq \left(\frac{|L_1(x)| + \ldots + |L_N(x)|}{N} \right)^N = (g \circ C)^N.$$

Applying the distance function formulation of MCBT, we get

$$\min(f) \leq \min((g \circ C)^N) = \min(g \circ C)^N \leq \frac{2^N}{\mathrm{Vol}(C\Omega)} = \left(\frac{N!}{N^N} \right) D.$$

We record the preceding work as follows.

Theorem 9.4. *(Product of Linear Forms) Let $C = (c_{ij}) \in \mathrm{GL}_N(\mathbb{R})$, and for $1 \leq i \leq N$, put $L_i(x) = \sum_{j=1}^N c_{ij} x_j$. Then there is $x = (x_1, \ldots, x_N) \in (\mathbb{Z}^N)^\bullet$ such that*

$$f(x) = |L_1(x) \cdots L_N(x)| \leq \left(\frac{N!}{N^N} \right) |\det C|. \tag{18}$$

Clearly (9.6) is an improvement over (9.5) for every $N > 1$, and a substantial improvement for large N.

Example: Taking $N = 2$ in Theorem 9.4, we find that there are integers x and y, not both zero, such that $|L_1(x,y)L_2(x,y)| \leq \frac{D}{2}$. In this case the problem is precisely that of the homogeneous minimum of an *indefinite binary quadratic form*. We will investigate this in the next section and find that the sharp bound is $|L_1 L_2| \leq \frac{D}{\sqrt{5}}$.

Example: Taking $N = 3$ in Theorem 9.4, we find that there are integers x, y, z, not all zero, such that $|L_1(x,y,z)L_2(x,y,z)L_3(x,y,z)| \leq \frac{2D}{9}$. In this case finding the best bound is already a piece of 20th century mathematics: it is a result of Davenport that one can have $f = |L_1 L_2 L_3| \leq \frac{D}{7}$, with equality iff f is equivalent under a $\mathbb{Z}$-linear change of variables to a scalar multiple of the cubic form

$$(x + 2\cos\frac{2\pi}{7}y + 2\cos\frac{4\pi}{7}z)(x + 2\cos\frac{4\pi}{7}y + 2\cos\frac{6\pi}{7}z)(x + 2\cos\frac{6\pi}{7}y + 2\cos\frac{8\pi}{7}z).$$

9.4. **Positive Definite Quadratic Forms.**

Lemma 9.5. *Let* $q(t) = q(t_1, \ldots, t_N) = \sum_{1 \le i \le j \le N} a_{ij} t_i t_j \in \mathbb{R}[t_1, \ldots, t_n]$ *be a real quadratic form which is* ***positive definite****:* $q(x) > 0$ *for all* $x \in (R^N)^\bullet$. *Let* $\operatorname{disc} q$ *be the determinant of the matrix* M_q *with* (i,j) *entry* a_{ij} *if* $i = j$ *and* $\frac{a_{ij}}{2}$ *if* $i \neq j$. *Put*

$$\Omega_R = \{x \in \mathbb{R}^N \mid q(x) \le R^2\}.$$

Then

$$\operatorname{Vol} \Omega_R = \frac{V_N}{(\operatorname{disc} q)^{\frac{1}{2}}} R^N. \tag{19}$$

Proof. Step 1: Suppose $q = q_0 = t_1^2 + \ldots + t_N^2$. Then Ω_R is nothing else than the R-ball with respect to the L^2-norm, so its volume is $R^N V_N$. Note that $\operatorname{disc} q_0 = 1$, so this verifies our claim in this case. Note
Step 2: Suppose q is diagonal: $q = a_1 t_1^2 + \ldots + a_N t_N^2$, so $\operatorname{disc} q = a_1 \cdots a_N$. Then

$$\Omega_R = \{x \in \mathbb{R}^N \mid a_1 x_1^2 + \ldots + a_N x_N^2 \le R^2\} = \{x \mid \left(\frac{x_1}{\sqrt{a_1}}\right)^2 + \ldots + \left(\frac{x_N}{\sqrt{a_N}}\right)^2 \le R^2\}.$$

Let $y = (y_1, \ldots, y_N) = (\frac{x_1}{\sqrt{a_1}}, \ldots, \frac{x_N}{\sqrt{a_N}})$. Then $q(x) = q_0(y)$, so making this linear change of variables transforms $\Omega_R(q)$ to $\Omega_R(q_0)$. Thus $\operatorname{Vol}(\Omega_R(q))$ is equal to $\operatorname{Vol}(\Omega_R(q_0))$ times the determinant of the linear transformation $x \mapsto y$, i.e.,

$$\operatorname{Vol}(\Omega_R(q)) = \frac{\operatorname{Vol}(\Omega_R(q_0))}{\sqrt{a_1 \cdots a_n}} = \frac{V_N}{(\operatorname{disc}(q))^{\frac{1}{2}}} R^N.$$

Step 3: By the Spectral Theorem, any real symmetric matrix may be orthogonally diagonalized: there is an orthogonal matrix P such that $q(Px)$ is diagonal. Orthogonal transformations leave the volume unchanged. On the matrix side, $M_{q(y)} = P^T M_q P$, so $\operatorname{disc} q(y) = \det P^T M_q P = (\det P)^2 \det M_q = \operatorname{disc} q$. Thus we have reduced to the diagonal case of Step 2. □

Theorem 9.6. *(Minkowski, 1891) Let* $q(t_1, \ldots, t_N) = \sum_{1 \le i \le j \le N} a_{ij} t_i t_j \in \mathbb{R}[t_1, \ldots, t_n]$ *be a positive definite real quadratic form. Let* $\Lambda \subset \mathbb{R}^n$ *be a lattice. There exists* $v \in \Lambda^\bullet$ *such that*

$$q(v) \le \frac{4(\operatorname{disc} q)^{\frac{1}{N}}}{V_N^{\frac{2}{N}}} (\operatorname{Covol} \Lambda)^{\frac{2}{N}}.$$

Proof. Note that $f(x) = \sqrt{q(x)}$ is a symmetric convex distance function in the sense of §2.4, with associated closed symmetric convex bodies the ellipsoids

$$\Omega_R = \{x \in \mathbb{R}^N \mid q(x) \le R^2\}.$$

By Minkowski's Convex Body Theorem (Mark II) we have a nonzero lattice point in $\{v \in \mathbb{R}^N \mid q(v) \le R^2\}$ when

$$\frac{V_N}{(\operatorname{disc} q)^{\frac{1}{2}}} R^N = \operatorname{Vol} \Omega_R = 2^N \operatorname{Covol} \Lambda,$$

i.e., if $R^2 = \frac{4(\operatorname{disc} q)^{\frac{1}{N}}}{V_N^{\frac{2}{N}}} (\operatorname{Covol} \Lambda)^{\frac{2}{N}}$. So there is $v \in \Lambda^\bullet$ with

$$q(v) \le R^2 = \frac{4(\operatorname{disc} q)^{\frac{1}{N}}}{V_N^{\frac{2}{N}}} (\operatorname{Covol} \Lambda)^{\frac{2}{N}}.$$

□

So there is a constant c_N such that for all positive definite, real n-ary forms q,

$$\min_{x \in (\mathbb{Z}^N)^\bullet} q(x) \leq c_N |\operatorname{disc} q|^{\frac{1}{N}}.$$

This qualitative form of Theorem 9.6 was proven in 1850 by Hermite. Hermite's arguments gave an explicit value of c_N, but it is much worse than the constant

$$M_N = 4V_N^{-\frac{2}{N}}$$

given by Minkowski's Theorem. Contemplating his improvement of Hermite's theorem Minkowski realized that his argument applied to much more general sets than level sets of positive definite quadratic forms – namely to symmetric convex bodies – and thus the Geometry of Numbers was born.

Table of Values of M_N:

$$M_2 = \frac{4}{\pi} = 1.2732395\ldots$$
$$M_3 = \left(\frac{6}{\pi}\right)^{\frac{2}{3}} = 1.5393389\ldots$$
$$M_4 = \frac{4\sqrt{2}}{\pi} = 1.8006326\ldots.$$
$$M_5 = 2.05845\ldots.$$
$$M_6 = 2.313629796\ldots$$
$$M_7 = 2.566728336\ldots$$
$$M_8 = 2.8181423672\ldots$$
$$M_9 = 3.068162\ldots$$
$$M_{10} = 3.3170068\ldots$$

We saw that the constant in Minkowski's Convex Body Theorem cannot be improved, but to see that it cannot be improved we took convex bodies associated to the L^∞-norm. It is reasonably to expect that when we restrict to ellipsoids the constant can be improved. This leads to the following key definition: for a function $f : \mathbb{R}^N \to \mathbb{R}^{\geq 0}$, we define its **homogeneous minimum**

$$m(f) = \inf_{x \in (\mathbb{Z}^N)^\bullet} f(x).$$

For any nondegenerate N-ary real form q, we define the **Hermite invariant**

$$\gamma(q) = \frac{m(q)}{|\operatorname{disc} q|^{\frac{1}{N}}}.$$

The next exercise addresses the sense in which $\gamma(q)$ is actually an "invariant" of q.

Exercise: Let q_1 and q_2 be two N-ary real quadratic forms.
a) Suppose that q_1 and q_2 are **integrally equivalent**: i.e., there exists $A \in \mathrm{GL}_N(\mathbb{Z})$ such that $q_1(At) = q_2(t)$. Show that $\gamma(q_1) = \gamma(q_2)$.
b) Suppose that q_1 and q_2 are **homothetic**: i.e., there exists $\alpha \in \mathbb{R}^{>0}$ such that $q_2 = \alpha q_1$. Show that $\gamma(q_1) = \gamma(q_2)$.
c) Let us say that q_1 and q_2 forms are **H-equivalent** if there exists $A \in \mathrm{GL}_N(\mathbb{Z})$ and $\alpha \in \mathbb{R}^{>0}$ such that $q_2(t) = \alpha q_1(At)$. By parts a) and b) above, H-equivalent

forms have the same Hermite invariant. Show that H-equivalence is actually an equivalence relation on (i) all nondegenerate n-ary real quadratic forms and (ii) all positive definite n-ary real quadratic forms.

For $N \in \mathbb{Z}^+$, define the **Hermite constant**

$$\gamma_N = \sup_q \gamma(q),$$

where q ranges over all *positive definite* N-ary quadratic forms.

Exercise: Show that $\gamma(q) = 1$ for *every* nondegenerate unary quadratic form q; in particular, $\gamma_1 = 1$.

Here is a very simple-minded way to give a lower bound on γ_N: find a positive definite *integral* quadratic form $q(t) \in \mathbb{Z}[t] = \mathbb{Z}[t_1, \ldots, t_n]$ which integrally represents 1. Then we must have $m(q) = 1$. An easy way to ensure an integral representation of 1 is to have $a_{11} = 1$: then $q(1, 0, \ldots, 0) = 1$. We try choose the form to have as small discriminant as possible: this gives a better bound on γ_N.

Example: For any $N \in \mathbb{Z}^+$, take $q_0 = x_1^2 + \ldots + x_N^2$. Then clearly $\min(q_0) = 1 = \operatorname{disc}(q_1)$, so $\gamma(q_0) = 1$. We deduce:

Proposition 9.7. *We have $\gamma_N \geq 1$ for all $N \in \mathbb{Z}^+$.*

Example: Let $N = 2$. Then, as above, we have everyone's favorite binary form $q_0 = x^2 + y^2$, with Hermite invariant 1. This isn't bad, but we can do better with the nondiagonal form $q_1(x, y) = x^2 + xy + y^2$. Again $m(q_1) = 1$, but now

$$\operatorname{disc} q_1 = \det \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix} = \frac{3}{4},$$

so $\gamma(q_1) = \frac{1}{\sqrt{\frac{3}{4}}} = \frac{2}{\sqrt{3}}$. Can we do even better than this? Not by this method: the general integral binary quadratic form $q(x, y) = ax^2 + bxy + cy^2$ has associated matrix $\begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}$ and thus discriminant $ac - \frac{b^2}{4} = \frac{4ac - b^2}{4}$. The numerator is a positive integer congruent to 0 or 3 mod 4, so the smallest it can take is 3 and thus $\operatorname{disc} q \geq \frac{3}{4}$. So let's record what we have so far:

$$1.15470\ldots = \frac{2}{\sqrt{3}} \leq \gamma_2 \leq M_2 = \frac{4}{\pi} = 1.2732\ldots$$

It turns out that $\gamma_2 = \frac{2}{\sqrt{3}}$. In fact even more is true: if a positive definite binary quadratic form q has Hermite invariant $\frac{2}{\sqrt{3}}$, then q is H-equivalent to $x^2 + xy + y^2$. This is a theorem of **Lagrange** that we will prove later on in this section.

Example: Let $N = 3$. Again, we can do better than $q_0 = x^2 + y^2 + z^2$ simply by taking all cross-terms with coefficient 1: $q_2 = x^2 + y^2 + z^2 + xy + xz + yz$. Then $m(q_2) = 1$ and $\det(q_2) = \frac{1}{2}$, so $\gamma(q_2) = 2^{\frac{1}{3}}$. Thus

$$1.25992\ldots = 2^{\frac{1}{3}} \leq \gamma_3 \leq M_3 = \left(\frac{6}{\pi}\right)^{\frac{2}{3}} = 1.5393389\ldots.$$

It is a theorem of **Gauss** that $\gamma_3 = 2^{\frac{1}{3}}$, and moreover a positive definite ternary quadratic form has Hermite invariant $2^{\frac{1}{3}}$ iff it is H-equivalent to q_2. We will give a proof later in these notes when we discuss reduction theory.

Example: Let $N = 4$. Based on the previous examples, it is natural to consider $q_3 = x^2 + y^2 + z^2 + w^2 + xy + xz + xw + yz + yw + zw$, for which we have $\gamma(q_3) = (5/16)^{\frac{-1}{4}} = 1.33748\ldots$. In fact though, if we look even through forms with small integer coefficients, we soon find $q_4 = x^2 + xz + y^2 - yz + z^2 - zw + w^2$, with $\gamma(q_4) = \sqrt{2} = 1.41421\ldots$. Thus

$$1.41421\ldots = \sqrt{2} \leq \gamma_4 \leq M_4 = \frac{4\sqrt{2}}{\pi} = 1.8006326\ldots.$$

It is a theorem of **Korkine-Zolotarev** that $\gamma_4 = \sqrt{2}$, and moreover a positive definite quaternary form q has $\gamma(q) = \sqrt{2}$ iff q is H-equivalent to q_4. Later we will follow a method of Mordell which enables us to derive γ_4 from knowledge of γ_3.

Exercise: Use the Minkowski-Hlawka Theorem to give an explicit lower bound on γ_N for all $N \geq 2$. (When I did this calculation, I got $\gamma_N \geq (2V_N\zeta(N))^{\frac{-2}{N}}$, but I didn't check this against anything in the literature, so don't take my word for it!)

Theorem 9.8. *There are positive constants C_1, C_2 such that for all $N \in \mathbb{Z}^+$,*

$$C_1 \leq \frac{\gamma_N}{N} \leq C_2.$$

Exercise: Use the previous exercise to prove Theorem 9.8. (Hint: given the previous exercise, much of the additional work is undergraduate advanced calculus involving the Γ function. For instance, Stirling's formula should be helpful here.)

9.5. **Binary Quadratic Forms.**

In this section we consider binary quadratic forms over the real numbers. Throughout, we let $A, B, C \in \mathbb{R}$ and consider

$$Q(x, y) = Ax^2 + Bxy + Cy^2.$$

We define

$$\Delta = \Delta(Q) = B^2 - 4AC.$$

Note that $\Delta(Q) = -4(\operatorname{disc} Q)$, where $\operatorname{disc} Q = AC - \frac{B^2}{4}$ is the determinant of the associated symmetric matrix

$$\begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix}.$$

We say the binary form Q is **degenerate** if $\Delta = 0$ and **nondegenerate** if $\Delta \neq 0$.

Lemma 9.9. *Let K be a field of characteristic different from 2, and let $Q(x, y) = Ax^2 + Bxy + Cy^2 \in K[x, y]$ be a quadratic form. TFAE:*
(i) Q is degenerate: $B^2 = 4AC$.
(ii) There exists a linear form $L(x, y) = \alpha x + \beta y$ and $\gamma \in K^\times$ such that

$$Q(x, y) = \gamma L(x, y)^2.$$

Exercise:
a) Prove Lemma 9.9.
b) Use Lemma 9.9 to show: if $Q(x,y) \in \mathbb{R}[x,y]$ is a degenerate binary quadratic form, then for any $\epsilon > 0$ there exist $(x,y) \in (\mathbb{Z}^2)^\bullet$ such that $|Q(x,y)| < \epsilon$.

Recall that for any domain R, a quadratic form $Q(x) = q(x_1, \ldots, x_n) \in R[x_1, \ldots, x_n]$ is **isotropic** if there exists $x \in (R^n)^\bullet$ such that $Q(x) = 0$; otherwise Q is **anisotropic**.

Proposition 9.10. *Let R be a domain of characteristic different from 2 with fraction field K. Let $Q(x,y) = Ax^2 + Bxy + Cy^2 \in R[x,y]$ be a binary quadratic form. Consider the following conditions:*
(i_R) There exist $\alpha, \beta, \gamma, \delta \in R$ such that

$$Q(x,y) = (\alpha x + \beta y)(\gamma x + \delta y).$$

(i_K) There exist $\alpha, \beta, \gamma, \delta \in K$ such that

$$Q(x,y) = (\alpha x + \beta y)(\gamma x + \delta y).$$

(ii_R) $\Delta = B^2 - 4AC$ is a square in R.
(ii_K) $\Delta = B^2 - 4AC$ is a square in K.
(iii_R) $Q(x,y)$ is isotropic over R.
(iii_K) $Q(x,y)$ is isotropic over K.
a) We have (i_R) $\implies$ (i_K), (ii_R) $\implies$ (ii_K), and (iii_R) $\iff$ (iii_K).
b) We have (i_K) $\iff$ (ii_K) $\iff$ (iii_K).
c) If R is integrally closed, then (ii_R) $\iff$ (ii_K).
d) If R is a UFD, then (i_R) $\iff$ (i_K) and thus all six conditions are equivalent.

Exercise: a) Prove Proposition 9.10. (Suggestion: if you are not a fan of commutative algebra, just prove it when $R = K$ is a field and when $R = \mathbb{Z}$.)
b) If you are really a fan of commutative algebra, try to construct examples to show that the additional hypotheses in Proposition 9.10 c) and d) above are needed in the sense that there are some domains R and binary forms $Q(x,y)$ for which the equivalences do not hold.

Remark 5. *The equivalence (iiiR) $\iff$ (iiiK) holds for forms in any number of variables. The other results are very particular to binary forms, as the following exercise makes clear.*

Exercise: Let R be a domain of characteristic different from 2 with fraction field K. Let $q(x) = q(x_1, \ldots, x_n)$ be a quadratic form over R with $n \geq 3$.
a) Suppose that q is nondegenerate in the sense that its defining symmetric matrix has nonzero determinant. Show that q is irreducible as a polynomial in K.
b) For those who know some algebraic geometry: a quadratic form $q(x_1, \ldots, x_n)$ defines a **quadric**, i.e., a projective hypersurface

$$Q : q(x) = 0$$

in $\mathbb{P}^n_{/K}$. Still assuming $\operatorname{char} K \neq 2$, show that the quadratic Q is smooth and geometrically irreducible iff the quadratic form q is nondegenerate.
c) Show that the results of the previous parts break down in characteristic 2.

Let $Q(x,y)$, $Q'(x',y')$ be two binary real quadratic forms. We say that Q and Q' are $\mathrm{SL}_2(\mathbb{Z})$-**equivalent** – and write $Q \sim Q'$ – if there is $M \in \mathrm{SL}_2(\mathbb{Z})$ such that

$$\left[\begin{array}{c} x \\ y \end{array} \right] = M \left[\begin{array}{c} x' \\ y' \end{array} \right].$$

In terms of the defininining symmetric matrices A_Q and $A_{Q'}$ this is the usual congruence relation

$$A_Q = M^T A_{Q'} M$$

with $M \in \mathrm{SL}_2(\mathbb{Z})$. We also have the notion of $\mathrm{SL}_N(\mathbb{Z})$-equivalence for real quadratic forms $Q(x_1,\ldots,x_N)$ and $Q'(x'_1,\ldots,x'_N)$.

For any real quadratic form $Q(x) = Q(x_1,\ldots,x_N)$, we define the **homogeneous minimum** of Q on a lattice $\Lambda \subset \mathbb{R}^N$

$$\min(Q,\Lambda) = \inf_{v \in \Lambda^\bullet} |Q(v)|.$$

We abbreviate $\min(Q,\mathbb{Z}^N)$ to $\min Q$.

Lemma 9.11. *If $Q(x)$ and $Q'(x')$ are $\mathrm{GL}_N(\mathbb{Z})$-equivalent real quadratic forms, then:*
a) We have $|\operatorname{disc} Q| = |\operatorname{disc} Q'|$.
b) For any lattice $\Lambda \subset \mathbb{R}^N$ we have $\min(Q,\Lambda) = \min(Q',\Lambda)$.

Exercise: Prove Lemma 9.11.

Now we can introduce the basic idea of **reduction** of real quadratic forms. Given a real quadratic form $Q(x) = Q(x_1,\ldots,x_N)$, we wish to exploit Lemma 9.11 to find $Q' \sim Q$ for which $m(Q',\Lambda)$ is easier to compute. It turns out in many cases there is a canonical "best" representative of the $\mathrm{SL}_N(\mathbb{Z})$-equivlaence class of Q, and the process of replacing Q by this "best representative" Q' is called **reduction**.

In general this is a big production, but when $N = 2$ one can *just do it.*

Lemma 9.12. *(Binary Reduction) Let $Q(x,y) = Ax^2 + Bxy + Cy^2$ be a nondegenerate real binary quadratic form. Let $(x_0,y_0) \in \mathbb{Z}$ be coprime integers such that $Q(x_0,y_0) = M \neq 0$. Then there are $b,c \in \mathbb{R}$ such that $Q \sim Mx^2 + bxy + cy^2$ with*

$$-|M| < b \leq |M|.$$

Proof. Step 1: Let x_0, y_0 be relatively prime integers such that $M = Q(x_0,y_0) \neq 0$. Choose $x_1, y_1 \in \mathbb{Z}$ such that $x_0 y_1 - x_1 y_0 = 1$. Then

$$x = x_0 x' + x_1 y',\ y = y_0 x' + y_1 y'$$

lies in $\mathrm{SL}_2(\mathbb{Z})$ and transforms $Q(x,y)$ into $Q'(x',y')$ with

$$A' = Ax_0^2 + 2Bx_0y_0 + Cy_0^2 = Q(x,y) = M.$$

Step 2: Let $n \in \mathbb{Z}$. Then

$$x' = x'' + ny'',\ y' = y''$$

lies in $\mathrm{SL}_2(\mathbb{Z})$ and transforms $Q'(x',y')$ into

$$Q''(x'',y'') = a(x'')^2 + bx''y'' + c(y'')^2 = M(x'')^2 + (B' + nM)x''y'' + c(y'')^2.$$

Since $M \neq 0$, we may choose n such that $-|M| < b \leq |M|$, qed. □

Theorem 9.13. *Let $Q(x,y) = Ax^2+Bxy+Cy^2$ be real and positive definite. Then $\gamma(Q) \leq \frac{2}{\sqrt{3}}$, with equality holding iff Q is H-equivalent to $q_2(x,y) = x^2+xy+y^2$.*

Proof. Since Q is positive definite, $\lim_{||(x,y)||\to\infty} Q(x,y) = \infty$. Thus Q attains a nonzero minimum m on the subset $(\mathbb{Z}^2)^\bullet$, say at (x_0,y_0). Necessarily x_0 and y_0 are relatively prime (otherwise dividing through by their gcd would give a smaller value). By Lemma 9.12, $Q \sim Q' = mx^2+bxy+cy^2$ with $-m < b \leq m$. Evaluating Q' at $(0,1)$ gives $c \geq m$, and thus

$$\operatorname{disc}(Q) = \operatorname{disc}(Q') = mc - \frac{b^2}{4} \geq m^2 - \frac{m^2}{4} = \frac{3m^2}{4},$$

so $m = \min Q \leq \sqrt{\frac{4\operatorname{disc}(Q)}{3}}$.

In the above analysis, equality holds iff $c = m$ and $b = \frac{m}{2}$, so $Q \sim Q' = m(x^2+xy+y^2)$. Since $\min(x^2+xy+y^2) = 1$ and $\operatorname{disc}(x^2+xy+y^2) = \frac{3}{4}$, $\min Q' = m = \sqrt{\frac{4\operatorname{disc}(Q')}{3}}$. □

Of course this yields immediately a result promised above.

Corollary 9.14. *We have $\gamma_2 = \frac{2}{\sqrt{3}}$.*

Theorem 9.15. *Let $Q(x,y) = Ax^2+Bxy+Cy^2$ be real and indefinite: $B^2-4AC > 0$. Then $\gamma(Q) \leq \frac{2}{\sqrt{5}}$, with equality holding iff Q is H-equivalent to x^2+xy-y^2.*

Proof. Let $m = \min Q = \inf_{(x,y)\in(\mathbb{Z}^2)^\bullet} |Q(x,y)|$. First note that we may have $m = 0$, but in this case the result is vacuously true. So we may assume that $m \neq 0$, and since $\min(-Q) = \min(Q)$ and $\operatorname{disc}(-Q) = \operatorname{disc}(Q)$, it is no loss of generality to assume $m > 0$. Unfortunately we may no longer assume that the infimum is attained, which slightly complicates things: nevertheless there exist relatively prime integers (x_0,y_0) with $M = Q(x_0,y_0) > 0$ and $m \leq M < 2m$.

By Lemma 9.12, $Q \sim Q'(x,y) = Mx^2+bxy+cy^2$ with $-M < b \leq M$. Note also $0 < \Delta(Q) = \Delta(Q') = b^2-4Mc$, so $b^2 > 4Mc$ For all $x,y \in \mathbb{Z}$, if $Q(x,y) < m$ then – since $m = \min Q$ – we have $Q(x,y) \leq -m$. Now

$$Q(0,1) = c < \frac{b^2}{4M} \leq \frac{M}{4} \leq \frac{m}{2} < m.$$

As above, this forces $c \leq -m < 0$ and thus $|c| \geq m$. We reason in this way again:
If $b \geq 0$, then $Q(1,-1) = M-b+c = M-|b|+c \leq M-m < m$.
If $b < 0$, then $Q(1,1) = M+b+c = M-|b|+c \leq M-m < m$.
So either way we have $M-|b|-|c| \leq -m$, or equivalently

$$|b| \geq M+m-|c|.$$

It follows that

$$\Delta = b^2+4M|c| \geq (M+m-|c|)^2+4M|c| = (M-m+|c|)^2+4Mm \geq (M-m+|c|)^2+4m^2.$$

Case 1: Suppose $M-m+|c| \geq m$. Then $\Delta \geq m^2+4m^2 = 5m^2$.
Case 2: Suppose $M-m+|c| < m$. Adding this to $M-|b|-|c| \leq -m$ gives

$$|b| > 2M-m \geq m$$

and thus

$$\Delta = b^2+4M|c| > m^2+4Mm \geq m^2+4m^2 = 5m^2.$$

This establishes part a). As for part b), equality occurs in the above iff $-c = M = m = |b|$, so $Q' = m(x^2 \pm xy - y^2)$. Note that $x^2 + xy - y^2 \sim x^2 - xy - y^2$ (we leave this as an exercise with the hint that $x^2 + xy - y^2$ is multiplicative and repesents -1) and $\operatorname{disc}(x^2 \pm xy - y^2) = 5$. This establishes part b). □

The following simple result will be useful later.

Proposition 9.16. *Let $Q(x, y) = Ax^2 + Bxy + Cy^2 \in \mathbb{Z}[x, y]$. Suppose $A > 0$ and $\Delta = B^2 - 4AC = -4$. Then Then $Q \sim_{\mathrm{SL}_2(\mathbb{Z})} x^2 + y^2$.*

Exercise: Prove Proposition 9.16. (Hint: first use the known value of γ_2 to show that Q integrally represents 1.)

9.6. **The Lattice Constant of a Star Body.**

Let $\Omega \subset \mathbb{R}^N$ be a star body; we *do not* assume that Ω is bounded. A lattice $\Lambda \subset \mathbb{R}^N$ is Ω-**admissible** if $\Lambda \cap \Omega^\circ = \{0\}$. We define the **lattice constant** $\Delta(\Omega)$ as follows:

$$\Delta(\Omega) = \inf\{\operatorname{Covol}\Lambda \mid \Lambda \text{ is } \Omega\text{-admissible}\}.$$

Since $\inf \varnothing = \infty$, this means that we have $\Delta(\Omega) = \infty$ iff there are no admissible Ω-lattices. This can certainly happen: e.g. take $\Omega = \mathbb{R}^N$.

A lattice Λ is Ω-**critical** if it is Ω-admissible and $\operatorname{Covol}\Lambda = \Delta(\Omega)$.

Exercise: Let $\Omega_1 \subset \Omega_2$ be star bodies. Show: $\Delta(\Omega_1) \leq \Delta(\Omega_2)$.

Exercise: a) Suppose Ω is bounded. Show: $\Delta(\Omega) < \infty$.
b) Give an example of an unbounded star body Ω with $\Delta(\Omega) < \infty$.
c) Let $f(x_1, \ldots, x_N) = |x_1 \cdots x_N|$. Show that f is a symmetric pseudo-distance function. Let $\Omega = f^{-1}([0, 1])$ be the corresponding star body. Are there any Ω-admissible lattices?

Exercise: a) Let Ω be a symmetric convex body. Show that Minkowski's Convex Body Theorem is equivalent to the inequality

$$\Delta(\Omega) \geq 2^{-N} \operatorname{Vol}\Omega.$$

b) Let Ω be a symmetric star body. Show that the Minkowski-Hlawka Theorem is equivalent to the inequality

$$\Delta(\Omega) \leq \frac{\operatorname{Vol}\Omega}{2\zeta(N)}.$$

For a star body Ω, let $v(\Omega)$ be the supremum of $\operatorname{Vol}\mathcal{B}$ as $\mathcal{B}$ ranges over convex bodies $\mathcal{B} \subset \Omega$.

Theorem 9.17. *For any star body Ω we have*

$$\Delta(\Omega) \geq \frac{v(\Omega)}{2^N} > 0.$$

Exercise: Prove it.

Let $r, s \in \mathbb{N}$ with $r + s = N$. Put

$$f_{r,s} = \sqrt{|\sum_{i=1}^{r} x_i^2 - \sum_{j=1}^{s} x_j^2|}.$$

This is a symmetric pseudo-distance function; put

$$\Omega_{r,s} = \Omega_{f_{r,s}} = f_{r,s}^{-1}([0,1])$$

and

$$\Gamma_{r,s} = \Delta(\Omega_{r,s}).$$

Exercise: a) Show: $\Gamma_{r,s} = \Gamma_{s,r}$.
Henceforth we may, and shall, assume $r \geq s$.
b) Suppose $r_1 + s_1 = r_2 + s_2$, $r_2 \geq r_1 \geq s_1$ and $r_s \geq s_2$. Show

$$\Gamma_{r_1,s_1} \geq \Gamma_{r_2,s_2}$$

and deduce

$$\Gamma_{r_1,s_1} \geq \Gamma_{r_1+s_1}.$$

Let $q = q(x_1, \ldots, x_n) \in \mathbb{R}[x_1, \ldots, x_n]$ be a nondegenerate quadratic form. We put

$$m(q) = \inf_{x \in (\mathbb{Z}^N)^\bullet} |q(x)|$$

and

$$\gamma(q) = \frac{m(q)}{|\operatorname{disc} q|^{\frac{1}{N}}}.$$

This generalizes the previously defined Hermite invariant of a positive form q. Now recall Sylvester's Theorem: if $q = q(x_1, \ldots, x_n) \in \mathbb{R}[x_1, \ldots, x_n]$ is a nondegenerate quadratic form, there are unique $r, s \in \mathbb{N}$ with $r + s = N$ such that there is $M \in \mathrm{GL}_n(\mathbb{R})$ such that

$$q(Mx) = f_{r,s}(x).$$

Let us say that q is of **type (r,s)**. We define

$$\gamma_{r,s} = \sup_{q \in \mathbb{R}[x_1,\ldots,x_n] \text{of type (r,s)}} \gamma(q).$$

Exercise: a) Show that

$$\gamma_{r,s} = \Gamma_{r,s}^{\frac{-2}{r+s}}. \tag{20}$$

b) Deduce

$$\gamma_{r,s} = \gamma_{s,r} \leq \gamma_{r+s}.$$

10. More on Hermite Constants

10.1. **Hermite's bound on the Hermite constant.**

Theorem 10.1. *(Hermite, 1850) Let $q(t) = q(t_1, \ldots, t_n) \in \mathbb{R}[t_1, \ldots, t_n]$ be an anisotropic quadratic form. Then*

$$m(q) \leq \left(\frac{4}{3}\right)^{\frac{n-1}{2}} |\operatorname{disc} q|^{\frac{1}{n}}.$$

Proof. ([G, Thm. 7.5]) We go by induction on n, the case $n = 1$ being trivial. Suppose the result holds for all forms of dimension less than n.

By Hermite's Lemma, any minimal vector extends to a basis, so by replacing q by an H-equivalent form $q(At)$, we may assume that the minimum occurs at $e_1 = (1, 0, \ldots, 0)$, and thus $m(q) = |a_{11}|$, where $q(t) = t^T M t$.

Let $A_\varphi : \mathbb{R}^n \to \mathbb{R}^n$ be the $\mathbb{R}$-linear map given by $\varphi(e_1) = e_1$ and for all $1 < j \leq n$, $\varphi(e_j) = e_j' = e_j - \frac{a_{1j}}{a_1} e_1$. Put $q'(t) = q(A_\varphi t)$. Then q' has matrix (a_{ij}') which is the direct sum of a_{11} with a block diagonal matrix $\frac{1}{a_{11}} C$, $c_{ij} = a_{11}a_{ij} - a_{ii}a_{1j}$. In other words, $q'(t) = a_{11}t_1^2 + q_2(t_2, \ldots, t_n)$, with $\operatorname{disc} q_2 = \frac{1}{a_{11}^{n-1}} \det C$. Since $\det A_\varphi = 1$, $\operatorname{disc} q' = \operatorname{disc} q = a_{11}^{2-n} \det C$, we have $\det C = a_{11}^{n-2} \operatorname{disc} q$.

Write $w = \sum_j \lambda_j e_j \in \mathbb{Z}^n$. Since for $j > 1$, $e_j = e_j' + \frac{a_{1j}}{a_{11}} e_1$, we have

$$w = \left(\lambda_1 + \frac{a_{12}}{a_{11}}\lambda_2 + \ldots + \frac{a_{1n}}{a_{11}}\lambda_n\right) e_1 + \lambda_2 e_2' + \ldots + \lambda_n e_n' = \gamma e_1 + z,$$

say. Suppose z is a minimal vector for q_2. By induction, we find

$$|q(z)| = |q_2(\lambda_2, \ldots, \lambda_n)| \leq \left(\frac{4}{3}\right)^{\frac{n-2}{2}} |\frac{1}{a_{11}}| |\det C|^{\frac{1}{n-1}} = \left(\frac{4}{3}\right)^{\frac{n-2}{2}} |a_{11}|^{\frac{-1}{n-1}} |\operatorname{disc} q|^{\frac{1}{n-1}}.$$

Having chosen $\lambda_2, \ldots, \lambda_n \in \mathbb{Z}$ to make z minimal, choose $\lambda_1 \in \mathbb{Z}$ so $|\gamma| \leq \frac{1}{2}$. Then

$$m(q) = |a_{11}| \leq |q(w)| \leq \gamma^2 a_{11} + q(z)$$

$$\leq \frac{|a_{11}|}{4} + \left(\frac{4}{3}\right)^{\frac{n-2}{2}} |a_{11}|^{\frac{-1}{n-1}} |\operatorname{disc} q|^{\frac{1}{n-1}},$$

so

$$|a_{11}|^{\frac{n}{n-1}} \leq \frac{|a_{11}|^{\frac{n}{n-1}}}{4} + \left(\frac{4}{3}\right)^{\frac{n-2}{2}} |\operatorname{disc} q|^{\frac{1}{n-1}}.$$

Thus

$$|a_{11}|^{\frac{n}{n-1}} \leq \left(\frac{4}{3}\right)^{\frac{n}{2}} |\operatorname{disc} q|^{\frac{1}{n-1}}$$

and finally

$$|a_{11}| \leq \left(\frac{4}{3}\right)^{\frac{n-1}{2}} |\operatorname{disc} q|^{\frac{1}{n}}.$$

□

Restricting to positive definite forms, Theorem 10.1 yields the upper bound

$$\gamma_n \leq H_n = \left(\frac{4}{3}\right)^{\frac{n-1}{2}}.$$

Explicitly,

$$\gamma_2 \leq H_2 = \sqrt{\frac{4}{3}} = 1.1547\ldots$$

$$\gamma_3 \leq H_3 = 4/3 = 1.333\ldots$$

$$\gamma_4 \leq H_4 = (4/3)^{3/2} = 1.5396\ldots$$

$$\gamma_5 \leq H_5 = \frac{16}{9} = 1.777\ldots$$

$$\gamma_6 \leq H_6 = (4/3)^{5/2} = 2.0528\ldots$$

$$\gamma_7 \leq H_7 = \frac{64}{27} = 2.37037\ldots$$
$$\gamma_8 \leq H_8 = (4/3)^{7/2} = 2.73706794\ldots$$
$$\gamma_9 \leq H_9 = \frac{256}{81} = 3.16049382716\ldots$$
$$\gamma_{10} \leq H_{10} = (4/3)^{9/2} = 3.6494239\ldots$$

Exercise:
a) By comparing the tables of values for M_n and H_n, check that $H_n < M_n$ for $2 \leq n \leq 8$, and $H_n > M_n$ for $n = 9, 10$.
b) Show that $H_n > M_n$ for all $n \geq 9$.

10.2. **The Known Hermite Constants.**

There are precisely nine positive integers n for which the exact value of γ_n is known, including $\gamma_1 = 1$. In the following result, when extremal forms are asserted to be unique, this means that they are unique up to H-equivalence (what else?).

Theorem 10.2. *(The Known Hermite Constants)*
a) (Lagrange) $\gamma_2 = \sqrt{\frac{4}{3}} = 1.1547\ldots$ *There is a unique extremal form,*

$$q_2(x, y) = x^2 + xy + y^2.$$

b) (Gauss) $\gamma_3 = 2^{\frac{1}{3}} = 1.25992\ldots$ *There is a unique extremal form,*

$$q_3(x, y, z) = x^2 + y^2 + z^2 + xy + xz + yz.$$

c) (Korkine-Zolotarev) $\gamma_4 = \sqrt{2} = 1.4142\ldots$ *There is a unique extremal form,*

$$q_4(x, y, z, w) = x^2 + y^2 + z^2 + w^2 + xz + xw + yz + yw + zw.$$

d) (Korkine-Zolotarev) $\gamma_5 = 8^{1/5} = 1.5157\ldots$. *There is a unique extremal form,*

$$q_5 = \ldots.$$

e) (Blichfeldt) $\gamma_6 = (\frac{64}{3})^{\frac{1}{6}} = 1.665366\ldots$. *There is a unique extremal form,*

$$q_6 = \ldots.$$

f) (Blichfeldt) $\gamma_7 = (64)^{\frac{1}{7}} = 1.811447\ldots$. *There is a unique extremal form,*

$$q_7 = \ldots.$$

g) (Blichfeldt) $\gamma_8 = 2$. *There is a unique extremal form,*

$$q_8 = \ldots.$$

h) (Cohn-Kumar) $\gamma_{24} = 4$. *There is a unique extremal form, the* ***Leech lattice****.*

Theorem 10.3. *We have:*
a) (Classical) $\gamma_{1,1} = \frac{2}{\sqrt{5}}$.
b) (Davenport) $\gamma_{1,2} = \gamma_{2,1} = (\frac{2}{3})^{\frac{1}{3}}$.
c) (Oppenheim)$\gamma_{2,2} = (\frac{2}{3})^{\frac{1}{2}}$, $\gamma_{3,1} = \gamma_{1,3} = (\frac{4}{7})^{\frac{1}{4}}$.
d) (Margulis) If $r, s \geq 1$ *and* $r + s \geq 5$, *then* $\gamma_{r,s} = 0$.

10.3. **Mordell's Inequality.**

In [Mo44], Mordell introduced a clever dual lattice technique that gives upper bounds on Hermite constants in the definite case. His method was refined by Oppenhiem [Op46] and applied to all quadratic forms by Cassels.

Theorem 10.4. *a) (Mordell) For all $n \geq 2$, we have*

$$\gamma_n \leq \gamma_{n-1}^{\frac{n-1}{n-2}}. \tag{21}$$

b) (Cassels) For $r, s \in \mathbb{Z}^+$, we have

$$\Gamma_{r,s}^{r+s-2} \geq \min \Gamma_{r-1,s}^{r+s}, \Gamma_{r,s-1}^{r+s}. \tag{22}$$

Proof. a) (Oppenheim [Op46]) It will be easier to compute with the quantity $L_n = \gamma_n^n$. Note that $L_n = \sup_f \frac{m(f)^n}{\operatorname{disc} f}$ as f ranges over all positive definite real n-ary quadratic forms. Since this quantity is scale invariant, we may restrict to forms f with $\operatorname{disc} f = 1$, and then $L_n = \sup_f m(f)^n$.

Let M_f be the defining symmetric matrix of f. Then its inverse matrix is also symmetric and positive definite and hence is the defining matrix of a quadratic form F, which we call the **adjugate form** of F. This process of taking adjugates gives an involution on the space of all positive definite n-ary quadratic forms: the adjugate of F is clearly f. Moreover, if f and g are $\mathrm{GL}_n(\mathbb{Z})$-equivalent, so are F and G, and conversely.

By Hermite's Lemma, f is $\mathrm{GL}_n(\mathbb{Z})$-equivalent to a form g such that $b_{11} = g(e_1) = m(g) = m(f)$. Let G be the adjugate form of g, and let $G'(t_2, \ldots, t_n) = G(0, t_2, \ldots, t_n)$. Then G' is an $(n-1)$-ary positive definite form with $\operatorname{disc} G' = b_{11}$. (This follows from the interpretation of the adjugate as a matrix of cofactors together with the fact that $\operatorname{adj} \operatorname{adj} A = A$ since $\det A = 1$.) Now

$$m(F) = m(G) \leq m(G') \leq \gamma_{n-1} (\operatorname{disc} G')^{\frac{1}{n-1}},$$

so

$$m(F)^{n-1} \leq L_{n-1} b_{11} = L_{n-1} m(f).$$

Applying this result with (F, f) in place of (f, F) gives

$$m(f)^{n-1} \leq L_{n-1} m(F).$$

Combining these last two inequalities gives

$$m(F)^{(n-1)^2} \leq L_{n-1}^{n-1} m(f)^{n-1} \leq L_{n-1}^n m(F),$$

so

$$M(F)^{(n-1)^2-1} \leq L_{n-1}^n$$

and thus

$$m(F)^n \leq \left(L_{n-1}^n\right)^{\frac{n}{(n-1)^2-1}} = L_{n-1}^{\frac{n}{n-2}}.$$

Since this holds for all F with $\operatorname{disc} F = 1$, we get

$$L_n \leq L_{n-1}^{\frac{n}{n-2}},$$

or equivalently

$$\gamma_n \leq \gamma_{n-1}^{\frac{n-1}{n-2}}.$$

b) For now, see [C, X.3.2]. □

As Oppenheim notes in [Op46], given the finiteness of γ_2 (whose easy proof was given in Theorem 9.13 above), this argument in fact gives an inductive proof of the finitness of γ_n for all $n \geq 3$.

Comparing with the known Hermite constants of §8.3, we see that Mordell's inequality is an equality in at least two cases:[9]

$$\sqrt{2} = \gamma_4 \leq \gamma_3^{\frac{3}{2}} = (2^{\frac{1}{3}})^{\frac{3}{2}} = \sqrt{2},$$
$$2 = \gamma_8 \leq \gamma_7^{\frac{7}{6}} = (64^{\frac{1}{7}})^{\frac{7}{6}} = 2.$$

10.4. **Computation of γ_3 and γ_4.**

Here we will give Mordell's elementary proof that $\gamma_3 = 2^{\frac{1}{3}}$ [Mor48]. Then, by Mordell's Inequality (Theorem 10.4a)) we have $\gamma_4 \leq \sqrt{2}$. Since for the form $q_4(x, y, z, w) = \dots$ we have $\gamma(q_4) = \sqrt{2}$, we deduce that $\gamma_4 = \sqrt{2}$.

Theorem 10.5. *(Gauss) Let $q(x, y, z)$ be a positive definite real quadratic form. Then $\gamma(q) \leq 2^{\frac{1}{3}}$, with equality iff q is H-equivalent to $q_3(x, y, z) = x^2 + y^2 + z^2 + xy + xz + yz$.*

Proof. ([Mor48]) Let $q(x, y, z)$ be positive definite with minimum $m(q) = 1$. We must show $\operatorname{disc} q \geq \frac{1}{2}$, with equality iff q is H-equivalent to q_3.

• By Hermite's Lemma (Lemma 1.16), by replacing q with an integrally equivalent form we can assume that the minimum value of 1 is taken at $(1, 0, 0)$ and thus the coefficient of x^2 is 1. We may then write q in the form

$$q = (x + \mu y + \nu z)^2 + by^2 + 2fyz + cz^2,$$

so $\operatorname{disc} q = bc - f^2$. Let m be the minimum of the positive definite binary form $q'(y, z) = by^2 + fyz + cz^2$. Applying Lemma 9.12, after a $\mathrm{SL}_2(\mathbb{Z})$ change of variables we can put q' in the form $my^2 + f'yz + c'z^2$ with $|f'| \leq m \leq c'$. By making the $\mathrm{GL}_2(\mathbb{Z})$ change of variables $(y, z) \mapsto (-y, z)$ if necessary, we may thus assume that (in our original notation) $0 \leq 2f \leq b \leq c$.

• By making a substitution

$$x = x' + py + qz, \ p, q \in \mathbb{Z},$$

we may – and shall – assume that $|\mu|, |\nu| \leq \frac{1}{2}$.

• We record some inequalities obtained by plugging in particular values and using that q is positive definite with minimum 1. Namely, plugging in

$$(x, y, z) = (0, 1, 0), \ (0, 0, 1), \ (\epsilon, 1, -1) \text{ for } \epsilon \in \mathbb{Z},$$

we get

$$b + \mu^2 \geq 1,$$
$$c + \nu^2 \geq 1,$$
$$b + c - 2f + (\epsilon + \mu - \nu)^2 \geq 1.$$

• Put $b = \beta + f$, $c = \gamma + f$, so that the following inequalities hold:

$$\beta, \gamma, f \geq 0,$$
$$|\mu|, |\nu| \leq \frac{1}{2},$$

[9] So far as I know, whether there are any further instances of equality is an **open question**.

$$\beta + f + \mu^2, \gamma + f + \nu^2 \geq 1,$$
$$\beta + \gamma + (\epsilon + \mu - \nu)^2 \geq 1.$$

We must show

$$\operatorname{disc} q = bc - f^2 = (\beta + f)(\gamma + f) - f^2 = \beta\gamma + f(\beta + \gamma) \geq \frac{1}{2}.$$

Fix $\epsilon \in \{-1, 0, 1\}$ such that $|\epsilon + \mu - \nu| \leq \frac{1}{2}$. Then

$$\beta + \gamma \geq 1 - (\epsilon + \mu - \nu)^2 \geq \frac{3}{4}.$$

Case 1: Suppose that either $f + \mu^2 \geq 1$ or $f + \nu^2 \geq 1$. Then $f \geq \frac{3}{4}$, so

$$\operatorname{disc} q \geq f(\beta + \gamma) \geq \frac{9}{16} > \frac{1}{2}.$$

Case 2: Suppose that $f + \mu^2, f + \nu^2 < 1$, so that $1 - f - \mu^2, 1 - f - \nu^2 > 0$. It follows that

$$\beta\gamma \geq (1 - f - \mu^2)(1 - f - \nu^2),$$

so

$$\operatorname{disc} q \geq \beta\gamma + f(\beta + \gamma) \geq (1 - f - \mu^2)(1 - f - \nu^2) + \left(1 - (\epsilon + \mu - \nu)^2\right) f.$$

Taking $\epsilon = 0$, we get

$$\begin{aligned} \operatorname{disc} q &\geq (1 - f - \mu^2)(1 - f - \nu^2) + f(1 - (\mu - \nu)^2) \\ &= (1 - \mu^2)(1 - \nu^2) + f(-1 + 2\mu\nu) + f^2 \\ &(1 - \mu^2)(1 - \nu^2) - (\mu\nu - \frac{1}{2})^2 + (f + \mu\nu - \frac{1}{2})^2 \\ = \frac{3}{4} - \mu^2 - \nu^2 + \mu\nu + (f + \mu\nu - \frac{1}{2})^2 &\geq \frac{3}{4} - \nu^2 - v^2 + \mu\nu. \end{aligned}$$

Case 2a: Suppose $\mu\nu \geq 0$. Then after replacing x by $-x$ if necessary, we may assume $\mu, \nu \geq 0$, and then

$$\mu^2 + \nu^2 - \mu\nu = \mu^2 - \nu(\mu - \nu) = \nu^2 - \mu(\nu - \mu) \leq \frac{1}{4}.$$

It follows that $\operatorname{disc} q \geq \frac{1}{2}$. Further, if equality holds then

$$\mu\nu(\mu - \nu) = f + \mu\nu - \frac{1}{2} = (\mu - \frac{1}{2})(\nu - \frac{1}{2}) = 0.$$

These imply that $(\mu, nu) \in \{(1/2, 1/2), (1/2, 0), (0, 1/2)\}$.

• If $\mu\nu \neq 0$, then $\mu = \nu = \frac{1}{2}$, $f = \frac{1}{4}$, $\beta = 1 - f - \mu^2 = \frac{1}{2} = \gamma$, $b = c = \frac{3}{4}$, so

$$q = (x + \frac{1}{2}y + \frac{1}{2}z)^2 + \frac{3}{4}y^2 + \frac{1}{2}yz + \frac{3}{4}z^2 = x^2 + y^2 + z^2 + xy + xz + yz = q_3.$$

• If $(\mu, \nu) = (\frac{1}{2}, 0)$, then $f = \frac{1}{2}$, $\beta = 1 - f - \mu^2 = \frac{1}{4}$, $\gamma = 1 - f - \nu^2 = \frac{1}{2}$, $b = \beta + f = \frac{3}{4}$, $c = \gamma + f = 1$, so

$$q = (x + \frac{1}{2}y)^2 + \frac{3}{4}y^2 + yz + z^2 = x^2 + y^2 + z^2 + xy + yz.$$

The change of variables $(x, y, z) = (X, Y + Z, -Z)$ takes us back to q_3.

• If $(\mu, nu) = (0, \frac{1}{2})$, then interchanging y and z takes us back to the previous case.

Case 2b: Suppose $\mu\nu < 0$ and $|\mu - \nu| \leq \frac{1}{2}$. Then we have

$$\operatorname{disc} q \geq \frac{3}{4} - \nu^2 - v^2 + \mu\nu \geq \frac{3}{4} - (\mu - \nu)^2 \geq \frac{1}{2},$$

and equality cannot arise since μ and ν have different signs. Case 2c: Suppose $\mu\nu < 0$ and $|\mu - \nu| > \frac{1}{2}$, and thus that $|\mu| + |\nu| > \frac{1}{2}$. Taking ϵ to be -1 if $\mu > 0$ and -1 if $\mu < 0$, we get

$$\begin{aligned}\operatorname{disc} q &\geq (1-\mu^2-f)(1-\nu^2-f) + (1-(1-|\mu|-|\nu|)^2)f \\ &= (1-\mu^2)(1-\nu^2) + (-2+2|\mu|+2|\nu|-2|\mu||\nu|)f + f^2 \\ &= (1-\mu^2)(1-\nu^2) - (1-|\mu|^2)(1-|\nu|)^2 + (f-(1-|\mu|)(1-|\nu|)^2 \\ &\geq 2(1-|\mu|)(1-|\nu|)(|\mu|+|\nu|).\end{aligned}$$

Put

$$|\mu| = \frac{1}{2} - m,\ |\nu| = \frac{1}{2} - n,$$

so

$$0 \leq m, n \leq \frac{1}{2},\ m+n < \frac{1}{2}.$$

Then

$$\begin{aligned}\operatorname{disc} q &\geq 2(\frac{1}{2}+m)(\frac{1}{2}+n)(1-m-n) \\ &\geq \frac{1}{2}(1+2m+2n)(1-m-n) \\ &= \frac{1}{2}(1+(m+n)-2(m+n)^2) \geq \frac{1}{2}.\end{aligned}$$

If equality holds then $m = n = 0$. Adjusting the sign on x so that $\mu > 0$, this gives $\mu = \frac{1}{2}$, $\nu = \frac{-1}{2}$, $f = \frac{1}{4}$, $\beta = \gamma = \frac{1}{2}$, $b = c = \frac{3}{4}$, so

$$q = (x + \frac{1}{2}y - \frac{1}{2}z)^2 + \frac{3}{4}y^2 + \frac{1}{2}yz + \frac{3}{4}z^2 = x^2 + y^2 + z^2 + xy - xz,$$

The change of variables $(x, y, z) = (-X, -Y, Z)$ gets us back to the form considered in Case 2a, hence to the critical form q_3. □

Theorem 10.6. *We have* $\gamma_4 = \sqrt{2}$.

Proof. By Theorem 10.5, $\gamma_3 = 2^{\frac{1}{3}}$. By Mordell's Inequality (Theorem 10.4), $\gamma_4 \leq \gamma_3^{\frac{3}{2}} = \sqrt{2}$. On the other hand, the form

$$q_4 = x^2 + y^2 + z^2 + w^2 + xz + xw + yz + yw + zw$$

has $m(q_4) = 1$ and $\operatorname{disc} q_4 = \frac{1}{4}$ hence

$$\gamma(q_4) = \frac{m(q_4)}{(\operatorname{disc} q_4)^{\frac{1}{4}}} = \sqrt{2}.$$

It follows that $\gamma_4 = \sqrt{2}$. □

Remark 6. *The form q_4 is, up to H-equivalence, the unique extremal form, but the approach we have taken does not yield this (as far as I can see).*

10.5. **Computation of $\gamma_{2,1}$ and $\gamma_{2,2}$.**

Theorem 10.7. *Let $q(x,y,z) \in \mathbb{R}[x,y,z]$ be an indefinite nondegenerate ternary quadratic form. Then $\gamma(q) \leq \left(\frac{2}{3}\right)^{\frac{1}{3}}$, with equality holding iff q is H-equivalent to $x_1^2 + x_1x_2 - x_2^2 - x_2x_3 + x_3^2$. In particular, $\gamma_{2,1} = \left(\frac{2}{3}\right)^{\frac{1}{3}}$.*

Proof. See [C, pp. 45-51]. □

Corollary 10.8. *We have $\Gamma_{2,2} = \frac{3}{2}$ and thus $\gamma_{2,2} = \sqrt{\frac{2}{3}}$.*

Proof. Combining Theorem 10.4b) and Theorem 10.7 we get

$$\Gamma_{2,2}^2 \geq \min \Gamma_{1,2}^4, \Gamma_{2,1}^4 = (\frac{3}{2})^2,$$

so $\Gamma_{2,2} \geq \frac{3}{2}$ and $\gamma_{2,2} \leq \sqrt{\frac{2}{3}}$. On the other hand, the form

$$q = x_1^2 + 2x_1x_2 - x_2^2 - x_3^2 - x_4^2 + 2x_1x_3 + x_1x_4 + x_2x_4 + x_3x_4$$

is anisotropic of discriminant $\frac{9}{4}$, so $\gamma(q) \geq (\frac{9}{4})^{\frac{-1}{4}} = \sqrt{\frac{2}{3}}$. □

11. Applications of GoN: Algebraic Number Theory

11.1. **Basic Setup.**

Let K be a number field of degree n, that is a field extension of $\mathbb{Q}$ which is n-dimensional as a $\mathbb{Q}$-vector space. Let $\mathbb{Z}_K$ be the ring of integers of K, i.e., the set of all elements of K satisfying a monic polynomial with integer coefficients. In other words, $\mathbb{Z}_K$ is the integral closure of $\mathbb{Z}$ in the finite dimensional field extension, and thus by [CA, §18.1] $\mathbb{Z}_K$ is a Dedekind domain: a Noetherian domain in which every nonzero prime ideal is maximal. Because $K/\mathbb{Q}$ is separable, $\mathbb{Z}_K$ is a finitely generated torsionfree $\mathbb{Z}$-module; since $\mathbb{Z}$ is a PID, $\mathbb{Z}_K$ is therefore isomorphic as a $\mathbb{Z}$-module to $\mathbb{Z}^n$. A choice of $x_1, \ldots, x_n \in \mathbb{Z}_K$ generating $\mathbb{Z}_K$ as a $\mathbb{Z}$-module is called an **integral basis**.

Given an n-tuple $x_1, \ldots, x_n$ of elements of K, we define the **discriminant** $D(x_1, \ldots, x_n)$ as the determinant of the $n \times n$ matrix with (i,j) entry $\mathrm{Tr}_{K/\mathbb{Q}}(x_ix_j)$. If $(x_1, \ldots, x_n) \in \mathbb{Z}_K$, then $D(x_1, \ldots, x_n) \in \mathbb{Z}$. Further, the discriminant scales under a linear change of variables as follows: if $(y_1, \ldots, y_n)^T = A(x_1, \ldots, x_n)^T$ for $A \in M_n(K)$, then

$$D(y_1, \ldots, y_n) = (\det A)^2 D(x_1, \ldots, x_n).$$

Since any two integral bases of $\mathbb{Z}_K$ are related via $A \in \mathrm{GL}_n(\mathbb{Z})$, this shows that any two integral bases have the same discriminant: we define this common value (an integer) to be the **discriminant** $d(K)$ of K.

Further, if $\sigma_1, \ldots, \sigma_n$ are the n field homomorphisms from $K \hookrightarrow \mathbb{C}$, then for $x_1, \ldots, x_n \in K$,

$$D(x_1, \ldots, x_n) = \det(\sigma_i(x_j))^2.$$

From this and Dedekind's linear independence of characters we get $d(K) \neq 0$.

11.2. **The Lattice Associated to an Ideal.**

Let $K/\mathbb{Q}$ be a number field of degree n. Concretely, $K \cong \mathbb{Q}[t]/(P(t))$, where $P(t)$ is an irreducible polynomial. It has r real roots and s pairs of complex roots, say, with $r + 2s = n$. Let us organize the corresponding field embeddings: $\sigma_1, \ldots, \sigma_r$ will be the real ones, and $\sigma_{r+1}, \ldots, \sigma_{r+s}$ will be pairwise nonconjugate complex embeddings. Define

$$\sigma : K \to \mathbb{R}^r \times \mathbb{C}^s,\ x \mapsto (\sigma_1(x), \ldots, \sigma_{r+s}(x)),$$

the **canonical embedding**. Since $r + 2s = n$, we may identify $\mathbb{R}^r \times \mathbb{C}^s$ with $\mathbb{R}^n$.

Proposition 11.1. *a) Let $M \subset K$ be a free $\mathbb{Z}$-submodule with basis $x_1, \ldots, x_n$. Then $\sigma(M)$ is a lattice in $\mathbb{R}^n$, with covolume*

$$\operatorname{Covol} M = 2^{-s} |\det_{1 \le i,j \le n} \sigma_i(x_j)|.$$

b) Let $\mathfrak{a}$ be a nonzero integral ideal of $\mathbb{Z}_K$. Then $\sigma(\mathfrak{a})$ is a lattice, with covolume

$$\operatorname{Covol} \mathfrak{a} = 2^{-s} \sqrt{|d(K)|} |\mathfrak{a}|$$

Proof. a) The image of x_i under σ with respect to the canonical basis of $\mathbb{R}^n$ is

$$v_i = (\sigma_1(x_i), \ldots, \sigma_r(x_i), \Re\sigma_{r+1}(x_i), \Im\sigma_{r+1}(x_i), \ldots, \Re\sigma_{r+s}(x_i), \Im\sigma_{r+s}(x_i)).$$

Clearly then $\sigma(M) = \langle v_1, \ldots, v_n \rangle$. To check that $\sigma(M)$ is a lattice we need to show that $v_1, \ldots, v_n$ are $\mathbb{R}$-linearly independent, and to compute the volume we need to compute the absolute value of the determinant of the matrix whose ith column is v_i. Thus the second task will imply the first as long as we get a nonzero volume. In fact it is easy to see that the determinant of the matrix is $(2i)^{-s} \det(\sigma_j(x_i))$, which has absolute value $2^{-s} |\det \sigma_i(x_j)|$. Since $x_1, \ldots, x_n$ form a K-base over $\mathbb{Q}$, the determinant is nonzero.
b) If we chose $M = \mathbb{Z}_K$, then since $(\det \sigma_i(x_j))^2 = d(K)$, the result follows in this case. In general, if $\mathfrak{a}$ is an ideal of R, then $|\mathfrak{a}| = \#\mathbb{Z}_K/\mathfrak{a}$. Thus $\sigma(\mathfrak{a})$ is an index $|\mathfrak{a}|$ sublattice of $\sigma(\mathbb{Z}_K)$, hence $\operatorname{Covol} \sigma(\mathfrak{a}) = |\mathfrak{a}| \operatorname{Covol} \sigma(\mathbb{Z}_K)$. The result follows. □

11.3. **A Standard Volume Calculation.**

Proposition 11.2. *Let $r, s \in \mathbb{N}$, $n = r + 2s$, $t \in \mathbb{R}$, and let*

$$B_t = \{(y_1, \ldots, y_r, z_1, \ldots, z_s) \in \mathbb{R}^r \times \mathbb{C}^s \mid \sum_{i=1}^{r} |y_i| + 2 \sum_{j=1}^{s} |z_j| \le t\}.$$

Then for all $t \ge 0$,

$$\operatorname{Vol} B_t = 2^r \left(\frac{\pi}{2}\right)^s \frac{t^n}{n!}.$$

Proof. ... □

As long as we are hiding – I mean, keeping – our volume calculations in a separate section, here is one more.

Proposition 11.3. *Let $r, s \in \mathbb{N}$, $n = r + 2s$, $d \in \mathbb{R}$. In $\mathbb{R}^r \times \mathbb{C}^s \cong \mathbb{R}^n$, consider the following set:*
• *If $r > 0$, $B = (y_1, \ldots, y_r, z_1, \ldots, z_s) \in \mathbb{R}^n$ such that $|y_1| \le 2^{n-1} \left(\frac{\pi}{2}\right)^{-s} \sqrt{|d|}$, $|y_i| \le \frac{1}{2}$ for $2 \le i \le r$, and $|z_j| \le \frac{1}{2}$ for $1 \le j \le s$.*

• *if* $r = 0$, $B = (y_1, \ldots, y_r, z_1, \ldots, z_s) \in \mathbb{R}^n$ *such that* $|z_1 - \overline{z_1}| \leq 2^n \left(\frac{\pi}{2}\right)^{1-s} \sqrt{|d|}$, $|z_1 + \overline{z_1}| \leq \frac{1}{2}$ *and* $|z_j| \leq \frac{1}{2}$ *for* $2 \leq j \leq s$. *Then*

$$\operatorname{Vol} B = 2^{n-s}\sqrt{|d|}.$$

Exercise: Prove Proposition 11.3.

11.4. **Finiteness of the Class Group.**

For a number field K of degree $n = r + 2s$, we define the **Minkowski constant**

$$M(K) = \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n} |d(K)|^{\frac{1}{2}}.$$

Theorem 11.4. *Let* $\mathfrak{a}$ *be a nonzero integral ideal of* $\mathbb{Z}_K$. *Then* $\mathfrak{a}$ *contains a nonzero element* x *such that*

$$|N_{K/\mathbb{Q}}(x)| \leq M(K)N(\mathfrak{a}).$$

Proof. Let $\sigma : K \to \mathbb{R}^r \times \mathbb{C}^s$ be the canonical embedding. Let $t \in \mathbb{R}^{>0}$, and as in Proposition 11.2 put

$$B_t = \{(y_1, \ldots, y_r, z_1, \ldots, z_s) \in \mathbb{R}^r \times \mathbb{C}^s \mid \sum_{i=1}^r |y_i| + 2\sum_{j=1}^s |z_j| \leq t\}.$$

B_t is a compact, symmetric convex body. Choose t such that

$$2^r \left(\frac{\pi}{2}\right)^s \frac{t^n}{n!} = \operatorname{Vol} B_t = 2^n \operatorname{Covol} \mathfrak{a} = 2^n 2^{-s}\sqrt{|d(K)|}N(\mathfrak{a}),$$

i.e., such that

$$t^n = 2^{n-r}\pi^{-s} n! \sqrt{|d(K)|}N(\mathfrak{a}).$$

By Minkowski's Convex Body Theorem, there is $x \in \mathfrak{a}^\bullet$ such that $\sigma(x) \in B_t$, so

$$|N_{K/\mathbb{Q}}(x)| = \prod_{i=1}^r |\sigma_i(x)| \prod_{j=r+1}^{r+s} |\sigma_j(x)|^2 \leq \left(\frac{1}{n}\sum_{i=1}^r |\sigma_i(x)| + \frac{2}{n}\sum_{j=r+1}^{r+s} |\sigma_j(x)|\right)^n \leq \frac{t^n}{n^n}$$

$$= \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n}\sqrt{|d(K)|}N(\mathfrak{a}) = M(K)N(\mathfrak{a});$$

the first inequality uses the AGM Inequality and the second the definition of B_t. □

Lemma 11.5. *Let* K *be a number field of degree* n, *and let* $r \in \mathbb{Z}^+$. *Then*

$$\#\{\mathfrak{a} \in \operatorname{Frac} \mathbb{Z}_K \mid \mathfrak{a} \supset \mathbb{Z}_K,\ [\mathfrak{a} : \mathbb{Z}_K] = r\} \leq 2^{r^n} < \infty.$$

Proof. If $\mathfrak{a} \supset \mathbb{Z}_K$ and $[\mathfrak{a} : \mathbb{Z}_K] = r$, then $r\mathfrak{a} \subset \mathbb{Z}_K$ and thus

$$\mathbb{Z}_K \subset \mathfrak{a} \subset \frac{1}{r}\mathbb{Z}_K.$$

Since $\frac{\frac{1}{r}\mathbb{Z}_K}{\mathbb{Z}_K} \cong (\mathbb{Z}/r\mathbb{Z})^n$, there are at most as many choices of $\mathfrak{a}$ as there are subsets of an r^n-element set (of course this is a ridiculously crude upper bound). □

Corollary 11.6. *Let* K *be a number field. There is a finite set* $I_1, \ldots, I_c$ *of fractional ideals of* $\mathbb{Z}_K$ *such that: for every nonzero ideal* $\mathfrak{a}$ *of* $\mathbb{Z}_K$, *there is* $1 \leq i \leq c$ *and* $\alpha \in K^\times$ *such that* $\mathfrak{a} = \alpha I_i$.

Proof. Let $\mathfrak{a}$ be a nonzero ideal of $\mathbb{Z}_K$. By Theorem 11.4, there is a nonzero element $\alpha \in \mathfrak{a}$ such that

$$[\mathbb{Z}_K : \alpha\mathbb{Z}_K] = |N_{K/\mathbb{Q}}(\alpha)| \leq M(K)N(\mathfrak{a}) = M(K)[\mathbb{Z}_K : \mathfrak{a}],$$

and thus we have

$$\left[\frac{1}{\alpha}\mathfrak{a} : \mathbb{Z}_K\right] = [\mathfrak{a} : \alpha\mathbb{Z}_K] \leq M(K).$$

By Lemma 11.5, the set of fractional ideals containing $\mathbb{Z}_K$ with index at most $M(K)$ is finite, and we are done. □

11.5. **Non-maximal orders.**

Let K be a degree n number field. An **order** in K is a subring $\mathcal{O}$ of K for which there exists a $\mathbb{Z}$-basis $x_1, \ldots, x_n$ for $\mathcal{O}$ which is also a $\mathbb{Q}$-basis for K.

The most important example of an order is the ring of integers $\mathbb{Z}_K$. In fact this is the unique maximal order: that is, every order $\mathcal{O}$ is contained in $\mathbb{Z}_K$, necessarily (just by virtue of its $\mathbb{Z}$-module structure) of some finite index f. We omit the detailed proof, but here is the basic idea: since $\mathcal{O}$ is finitely generated as a $\mathbb{Z}$-module, every element satisfies a monic polynomial with $\mathbb{Z}$-coefficients, i.e., is *integral* over $\mathbb{Z}$. But $\mathbb{Z}_K$ is nothing else than the set of all elements of K which are integral over $\mathbb{Z}$.

The maximal order $\mathbb{Z}_K$ has an important property that any non-maximal order lacks: it is integrally closed in K and thus is a **Dedekind domain**. Dedekind domains are characterized among all integral domains by the following fact: for every nonzero fractional ideal $\mathfrak{a}$, there is a fractional ideal $\mathfrak{b}$ such that $\mathfrak{ab} = \mathbb{Z}_K$: in other words, $\operatorname{Frac}\mathbb{Z}_K$ forms a *group* under multiplication. Letting $\operatorname{Prin}\mathbb{Z}_K$ denote the subgroup of principal fractional ideals $\alpha\mathbb{Z}_K$, we may form the quotient

$$\operatorname{Pic}\mathbb{Z}_K = \operatorname{Frac}\mathbb{Z}_K / \operatorname{Prin}\mathbb{Z}_K,$$

the **ideal class group** of $\mathbb{Z}_K$.

Exercise: Show that an equivalent restatement of Corollary 11.6 is that for any number field K, $\operatorname{Pic}\mathbb{Z}_K$ is a **finite abelian group**.

Now we consider the case of a non-maximal order $\mathcal{O}$. In this case $\operatorname{Frac}\mathcal{O}$ is a monoid under multiplication which is not a group: some elements have no inverse. We may still consider the quotient

$$C(\mathcal{O}) = \operatorname{Frac}\mathcal{O} / \operatorname{Prin}\mathcal{O}$$

of fractional ideals modulo principal ideals, the **ideal class monoid** of $\mathcal{O}$. Every element of $C(\mathcal{O})$ is represented by a nonzero integral ideal $\mathfrak{a}$ of R and two ideals $\mathfrak{a}, \mathfrak{b}$ determine the same element of $C(\mathcal{O})$ iff there are nonzero elements $\alpha, \beta \in R$ such that $\alpha\mathfrak{a} = \beta\mathfrak{b}$.

Here is the point: we claim that the proofs of the previous section have been crafted so as to be easily modified to establish the following result.

Theorem 11.7. *For any order $\mathcal{O}$ in a number field K, $C(\mathcal{O})$ is finite.*

Exercise: Let $\mathcal{O}$ be an order in a number field K, and put $f = [\mathbb{Z}_K : \mathcal{O}]$.
a) Show that for any nonzero ideal $\mathfrak{a}$ of $\mathcal{O}$, $\sigma(\mathfrak{a})$ is a lattice in $\mathbb{R}^n$ of covolume $2^{-s} f \sqrt{|d(K)|}[\mathcal{O} : \mathfrak{a}]$.
b) Show that for any nonzero ideal $\mathfrak{a}$, $\exists\ \alpha \in \mathfrak{a}^\bullet$ with $|N_{K/Q}(x)| \leq f M(K)[\mathcal{O} : \mathfrak{a}]$.
c) Prove Theorem 11.7.

Remark: For an arbitrary order $\mathcal{O}$, there is still a Picard group $\operatorname{Pic}\mathcal{O}$: it is the group of units of the monoid $C(\mathcal{O})$. In other words, to form $\operatorname{Pic}\mathcal{O}$ we consider only invertible fractional ideals (which, tautologically, form a group under multiplication) and quotient out by the subgroup of principal ideals. Because $\operatorname{Pic}\mathcal{O} \subset C(\mathcal{O})$, we immediately deduce the following result.

Corollary 11.8. *For any order $\mathcal{O}$ in a number field, $\operatorname{Pic}\mathcal{O}$ is a finite abelian group.*

One of the most important and deep problems in algebraic number theory is to understand the structure of the finite abelian group $\operatorname{Pic}\mathbb{Z}_K$ as K ranges over all number fields. One might think then that the study of Picard groups of nonmaximal orders $\mathcal{O} \subset \mathbb{Z}_K$ would be even worse, but in fact there is a natural surjection $\operatorname{Pic}\mathcal{O} \to \operatorname{Pic}\mathbb{Z}_K$ with an explicitly understood kernel.

11.6. **Other Finiteness Theorems.**

Theorem 11.9. *(Hermite-Minkowski) Let K be a number field of degree $n \geq 2$.*
a) We have $|d(K)| \geq \frac{\pi}{3}\left(\frac{3\pi}{4}\right)^{n-1}$.
b) In particular $|d(K)| > 1$.

Proof. a) By Corollary 11.6, there is a nonzero integral ideal $\mathfrak{b}$ with

$$1 \leq N(\mathfrak{b}) \leq M(K) = \left(\frac{4}{\pi}\right)^s \frac{n!}{n^n} |d(K)|^{\frac{1}{2}}.$$

Thus

$$|d(K)| \geq \left(\frac{\pi}{4}\right)^{2s} \frac{n^{2n}}{(n!)^2} \geq \left(\frac{\pi}{4}\right)^n \frac{n^{2n}}{(n!)^2} = a_n,$$

say. Note $a_2 = \frac{\pi^2}{4}$ and $\frac{a_{n+1}}{a_n} = \frac{\pi}{4}(1 + \frac{1}{n})^{2n} \geq \frac{3\pi}{4}$ by the binomial theorem. Thus for $n \geq 2$,

$$|d(K)| \geq \frac{\pi^2}{4}\left(\frac{3\pi}{4}\right)^{n-2} = \frac{\pi}{3}\left(\frac{3\pi}{4}\right)^{n-1}.$$

b) If $n \geq 2$, $|d(K)| \geq \frac{\pi}{3} \cdot \frac{3\pi}{4} = \frac{\pi^2}{4} > 1$. □

Remark: Actually the proof shows that if $n > 1$, $|d(K)| > 2$! Perhaps this result is not more widely advertised because of a theorem of Stickelberger that the discriminant of any number field must be congruent to 0 or 1 modulo 4.

Theorem 11.10. *(Hermite) There are up to isomorphism only finitely many number fields with a given discriminant $d \in \mathbb{Z}$.*

Proof. By Theorem 11.9, it suffices to show that for any fixed $r, s \in \mathbb{N}$, there are only finitely many number fields with r real places, s complex places, degree $n = r + 2s$ and discriminant d. Let $B \subset \mathbb{R}^r \times \mathbb{C}^s$ be as defined in Proposition 11.3. B is compact, convex and centrally symmetric, with $\operatorname{Vol} B = 2^{n-s}|\sqrt{d}|$. Applying Proposition 11.1 with $\mathfrak{a} = \mathbb{Z}_K$ and Minkowski's Convex Body Theorem, we get $x \in \mathbb{Z}_K^\bullet$ such that $\sigma(x) \in B$. We claim x is a primitive element of K, i.e., $K = \mathbb{Q}(x)$.

Suppose first that $r > 0$, so $|\sigma_i(x)| \leq \frac{1}{2}$ for $i \neq 1$. Since $|N(x)| = \prod_{i=1}^n |\sigma_i(x)| \in \mathbb{Z}^+$, we have $|\sigma_1(x)| \neq 1$ and thus $\sigma_1(x) \neq \sigma_i(x)$ for some $i > 1$. It follows that x is primitive. Similarly, if $r = 0$, then $|\sigma_1(x)| = |\overline{\sigma_1(x)}| \geq 1$, so $\sigma_1(x) \neq \sigma_j(x)$ when $\sigma_j \neq \sigma_1, \overline{\sigma_1}$. It follows that the real part of $\sigma_1(x)$ is at most $\frac{1}{4}$ in absolute value. Since $|\sigma_1(x)| \geq 1$, this implies $\sigma_1(x)$ is not real, and thus $\sigma_1(x) \neq \overline{\sigma_1(x)}$ and thus x is primitive.

The inequalities defining B show that all the conjugates $\sigma_i(x)$ are bounded, hence coefficients of the minimal polynomial of x, being elementary symmetric functions in the $\sigma_i(x)$'s, are also bounded, and this gives finitely many choices for x and thus finitely many choices for K. □

11.7. **The Dirichlet Unit Theorem.**

Let K be a number field. We wish to study the structure of the unit group $\mathbb{Z}_K^\times$.

Lemma 11.11. *For $x \in \mathbb{Z}_K$, the following are equivalent:*
(i) $x \in \mathbb{Z}_K^\times$.
(ii) $|N(x)| = 1$.

Proof. If $x \in \mathbb{Z}_K^\times$, there is $y \in \mathbb{Z}_K^\times$ such that $xy = 1$, and then $1 = |N(1)| = |N(xy)| = |N(x)|N(y)|$. Since $|N(x)|, |N(y)| \in \mathbb{Z}^+$, this forces $|N(x)| = 1$. Conversely, if $|N(x)| = 1$, the minimal polynomial of x over $\mathbb{Q}$ is $x^n + a_{n-1}x^{n-1} + \ldots + a_1 x \pm 1 = 0$, so $x \cdot (x^{n-1} + a_{n-1}x^{n-2} + \ldots + a_1) = \pm 1$, so $x \in \mathbb{Z}_K^\times$. □

Theorem 11.12. *Let K be a number field of degree $n = r + 2s$. Then $\mathbb{Z}_K^\times$ is a finitely generated abelian group, with free rank $r + s - 1$ and torsion subgroup the group $\mu(K)$ of roots of unity in K.*

Proof. Step 0: We define a homomorphism $L : \mathbb{Z}_K^\times \to \mathbb{R}^{r_1+r_2}$, a variant of the canonical embedding:

$$L : x \mapsto (\log|\sigma_1(x)|, \ldots, \log|\sigma_{r+s}(x)|).$$

Step 1: We claim that for any compact subset $B \subset \mathbb{R}^{r_1+r_2}$, $B' = L^{-1}(B)$ is finite. Because B is bounded, there is $\alpha > 1$ such that for all $x \in B'$, $\frac{1}{\alpha} \leq |\sigma_i(x)| \leq \alpha$. From this it follows that the elementary symmetric functions of the $\sigma_i(x)$'s are bounded; because they take integer values, their values are therefore restricted to lie in a finite set. It follows that there are only finitely many possible characteristic polynomials for $x \in B'$ and hence only finitely many possible values for such x.
Step 2: It follows from Step 1 that $L^{-1}(0) = \operatorname{Ker} L$ is finite. In particular, each element of $\operatorname{Ker} L$ has finite order, i.e., is a root of unity. Conversely, every root of unity in K is a unit of $\mathbb{Z}_K$ lying in the kernel of L, so $\operatorname{Ker} L = \mu(K)$, a finite group.
Step 3: It also follows from Step 1 that $L(\mathbb{Z}_K^\times)$ is a discrete subgroup of $\mathbb{R}^{r+s}$, hence free abelian of rank at most $r + s$. Moreover, for $x \in \mathbb{Z}_K^\times$, by Lemma 11.11 we have

$$\pm 1 = N(x) = \prod_{i=1}^n \sigma_i(x) = \prod_{i=1}^r \sigma_i(x) \prod_{j=r+1}^{r+s} \sigma_j(x)\overline{\sigma_j(x)},$$

hence $L(x)$ lies in the hyperplane

$$W : \sum_{i=1}^r y_i + 2 \sum_{j=r+1}^{r+s} = 0.$$

Thus $L(\mathbb{Z}_K^\times) \subset W \cong \mathbb{R}^{r+s-1}$, so is free abelian of rank at most $r+s-1$.
Step 4: The last, most delicate part of the argument, is to show that $L(\mathbb{Z}_K^\times)$ has rank $r+s-1$. We show this by a duality argument: for any nonzero linear form $f : W \to \mathbb{R}$, we claim there exists $u \in \mathbb{Z}_K^\times$ such that $W(L(u)) \neq 0$. Notice that this shows that $\langle L(\mathbb{Z}_K^\times)\rangle_{\mathbb{R}} = W$, which implies $L(\mathbb{Z}_K^\times) \cong \mathbb{Z}^{r+s-1}$.

Put $N = r+s-1$. The projection of W onto $\mathbb{R}^N$ is an isomorphism, so we may write, for any $y = (y_1, \ldots, y_{N+1}) \in W$,

$$f(y) = c_1 y_1 + \ldots + c_N y_N,\ c_i \in \mathbb{R}.$$

Fix a real number $\alpha \geq 2^n \left(\frac{1}{2\pi}\right)^s \sqrt{|d(K)|}$. For any $\lambda = (\lambda_1, \ldots, \lambda_N)$ with $\lambda_i > 0$ for all i, take $\lambda_{N+1} > 0$ such that

$$\prod_{i=1}^{r} \lambda_i \prod_{j=r+1}^{r+s} \lambda_j^2 = \alpha.$$

In $\mathbb{R}^r \times \mathbb{C}^s$, the set B of elements $(y_1, \ldots, y_r, z_1, \ldots, z_s)$ with $|y_i| \leq \lambda_i$ and $|z_j| \leq \lambda_{r+j}$ is a compact, symmetric convex set of volume

$$\prod_{i=1}^{r} 2\lambda_i \prod_{j=r+1}^{r+s} \pi\lambda_j^2 = 2^r \pi^s \alpha \geq 2^{n-s} |\sqrt{d(K)}|.$$

By MCBT and Proposition 11.1 there is $x_\lambda \in \mathbb{Z}_K^\bullet$ such that $\sigma(x_\lambda) \in B$. Thus $|\sigma_i(x_\lambda)| \leq \lambda_i$ for all i (for $j \geq s+1$, we put $\lambda_j = \lambda_{j-s}$). Thus

$$1 \leq |N(x_\lambda)| = \prod_{i=1}^{n} |\sigma_i(x_\lambda)| \leq \prod_{i=1}^{r} \lambda_i \prod_{j=r+1}^{r+s} \lambda_j^2 = \alpha.$$

Moreover, for all $1 \leq i \leq n$,

$$|\sigma_i(x_\lambda)| = |N(x_\lambda| \prod_{j \neq i} |\sigma_j(x_\lambda)|^{-1} \geq \prod_{j \neq i} \lambda_j^{-1} = \lambda_i \alpha^{-1}$$

so

$$\lambda_i \alpha^{-1} \leq |\sigma_i(x_\lambda)| \leq \lambda_i,$$

hence

$$0 \leq \log \lambda_i - \log |\sigma_i(x_\lambda)| \leq \log \alpha.$$

Applying the linear form f we get

$$|f(L(x_\lambda)) - \sum_{i=1}^{N} c_i \log \lambda_i| \leq \left(\sum_{i=1}^{N} |c_i|\right) \log \alpha = \gamma,$$

say. Let $\beta > \gamma$ be a constant, and for each $h \in \mathbb{Z}^+$, choose positive real numbers $\lambda_{1,h}, \ldots, \lambda_{N,h}$ such that $\sum_{i=1}^{N} c_i \log \lambda_{i,h} = 2\beta h$. Put $\lambda(h) = (\lambda_{1,h}, \ldots, \lambda_{N,h})$ and let $x_h = x_{\lambda(h)}$ be the corresponding element of $\mathbb{Z}_K^\bullet$. Then $|f(L(x_h)) - 2\beta h| < \beta$, so

$$(2h-1)\beta < f(L(x_h)) < (2h+1)\beta.$$

It follows that the $f(L(x_h))$ are all distinct. But since $|N(x_h)| \leq \alpha$, there are only finitely many principal ideals $x_h \mathbb{Z}_K$, so there exists $h \neq h'$ with $(x_h) = (x_{h'})$ and thus $x_h = u x_{h'}$ with $u \in \mathbb{Z}_K^\times$. Thus $f(L(u)) = f(L(x_h)) - f(L(x_{h'})) \neq 0$. □

11.8. **The Lattice Associated to an S-Integer Ring.**

Let K be a number field, and let S be a finite set of places of K containing all the Archimedean places. We write Σ_K for the set of all places of K, $\Sigma_{K,f} = \operatorname{MaxSpec} \mathbb{Z}_K$ for the set of all finite places, Σ_K^∞ for the set of all infinite places, and S_f for the set of finite places in S. Let $\mathbb{Z}_{K,S}$ denote the ring of S-integers of K, i.e., the set of all elements $x \in K$ such that $|x|_v \leq 1$ for all $v \notin S$. Most of the basic finiteness theorems of algebraic number theory are classically stated in terms of the rings $\mathbb{Z}_K$ but can be extended to the rings $\mathbb{Z}_{K,S}$. In fact, doing so is a matter of pure commutative algebra. For instance:

• Because $\operatorname{Pic} \mathbb{Z}_K$ is finite and $\mathbb{Z}_{K,S}$ is a ring intermediate between $\mathbb{Z}_K$ and its fraction field K – an **overring** – $\operatorname{Pic} \mathbb{Z}_{K,S}$ is also finite. More precisely,

$$\operatorname{Pic} \mathbb{Z}_{K,S} = \operatorname{Pic} \mathbb{Z}_K / \langle [\mathfrak{p}_v] \mid v \in S_f \rangle.$$

• Because $\mathbb{Z}_K^\times$ is finitely generated and S is finite, $\mathbb{Z}_{K,S}^\times$ is finitely generated. Because $\operatorname{Pic} \mathbb{Z}_K$ is torsion, we can be more precise:

$$\mathbb{Z}_{K,S}^\times \cong \mathbb{Z}_K^\times \oplus \mathbb{Z}^{\#S_f}.$$

Nevertheless, for certain nefarious purposes (not yet attained in these notes) it is desirable to extend the lattice perspective to rings of S-integers $\mathbb{Z}_{K,S}$. At first glance this seems unlikely: consider the simplest nontrivial case, $\mathbb{Z}[\frac{1}{p}]$ for some prime p. The additive group of this ring is *not* finitely generated, so it cannot be realized as a full lattice in any Euclidean space.

However, a more ambitious approach does work: namely, each ring $\mathbb{Z}_{K,S}$ can be naturally embedded in a locally compact topological ring in such a way that it is discrete and with compact quotient.

Coming back to our simple example $\mathbb{Z}[\frac{1}{p}]$, we need a ring in which the sequence $\frac{1}{p^n}$ does not converge to 0. For those who know about p-adic numbers – and those who do not may as well stop reading this section here: fair warning! – the natural choice to solve that problem is $\mathbb{Q}_p$. However, it is still not the case that $\mathbb{Z}[\frac{1}{p}]$ is discrete in $\mathbb{Q}_p$ because even its subring $\mathbb{Z}$ is not discrete in $\mathbb{Q}_p$: if so it would be closed, but its closure is $\mathbb{Z}_p$. A little thought shows that the desired embedding is

$$\Delta : \mathbb{Z}[\frac{1}{p}] \hookrightarrow \mathbb{Q}_p \times \mathbb{R}, \ x \mapsto (x, x).$$

Indeed, if we take on $\mathbb{Q}_p \times \mathbb{R}$ the metric which is the maximum of the standard metric on $\mathbb{R}$ and the standard (p-adic) metric on $\mathbb{Q}_p$, then $B(0,1) \cap \Delta(\mathbb{Z}[\frac{1}{p}]) = 0$. It is also easy to see that every element of the quotient $(\mathbb{Q}_p \times \mathbb{R})/\Delta(\mathbb{Z}[\frac{1}{p}])$ has a representative in $\mathbb{Z}_p \times [0,1]$, hence the quotient is compact. Thus it is reasonable to view $\Delta(\mathbb{Z}[\frac{1}{p}])$ as a **lattice** in the locally compact group $\mathbb{Q}_p \times \mathbb{R}$.

We now return to the general case of $\mathbb{Z}_{K,S}$. This time there is a natural diagonal embedding

$$\Delta : \mathbb{Z}_{K,S} \hookrightarrow K_S := \prod_{v \in S} K_v.$$

We claim that the image is discrete and the quotient $K_S/\Delta(\mathbb{Z}_{K,S})$ is compact. We will show this following some notes of B. Conrad.

Lemma 11.13. *Let K'/K be a finite extension of number fields, and let S be a finite set with $\Sigma_K^\infty \subset S \subset \Sigma_K$. Let S' be the set of places of K' lying over some place of S. Then the integral closure of $\mathbb{Z}_{K,S}$ in K' is $\mathbb{Z}_{K',S'}$.*

Proof. By [CA, Cor. 22.6], we may write $\mathbb{Z}_{K,S} = \mathbb{Z}_K[\frac{1}{a}]$ for some $a \in \mathbb{Z}_K^\bullet$, and we are reduced to showing that the integral closure of $\mathbb{Z}_K[\frac{1}{a}]$ in K' is $\mathbb{Z}_{K'}[\frac{1}{a}]$. This holds because integral closure commutes with localization [CA, Thm. 14.9]. □

Lemma 11.14. *Let G_1, G_2 be topological groups, with G_1 Hausdorff and G_2 quasi-compact. Let Γ be a discrete subgroup of $G_1 \times G_2$ such that $\pi_1 : \Gamma \to G_1$ is injective. Then $\pi_1(\Gamma)$ is discrete in G_1.*

Proof. If $\pi_1(\Gamma)$ is not discrete in G_1, there is a net $x_\bullet : I \to \pi_1(\Gamma) \setminus e_1$ such that $\pi_1(x_\bullet) \to e_1$. Now consider the net $\pi_2(x_\bullet)$ in G_2: since G_2 is quasi-compact, after passing to a subnet (we will not change the notation) we get $\pi_2(x_\bullet) \to g_2 \in G_2$. But then it follows that $x_\bullet \to (e_1, g_2)$. Since Γ is discrete in $G_1 \times G_2$, this net must be ultimately constant, and hence its image in G_1 is ultimately constant. Since G_1 is Hausdorff, the unique limit of the net $\pi(x_\bullet)$ is the eventually constant value which lies in $\pi_1(\Gamma) \setminus e_1$: contradiction. □

Theorem 11.15. *Let K be a number field, S a finite set of places of K containing all Archimedean places. Then the diagonal embedding*

$$\Delta : \mathbb{Z}_{K,S} \to K_S := \prod_{v \in S} K_v$$

has discrete and cocompact image.

Proof. Step 1: Let $S' \supset S$ be a finite subset of Σ_K. We show that if $\Delta(\mathbb{Z}_{K,S'})$ is discrete and cocompact in $K_{S'}$ then $\Delta(\mathbb{Z}_{K,S})$ is discrete and compact in K_S.

□

The locally compact group K_S carries a Haar measure, which we can take as the product of the Haar measures on the factors, and normalize the Haar measure on each non-Archimedean K_v by taking the one which gives the valuation ring unit measure.

Exercise: Compute the covolume of $\Delta(\mathbb{Z}_{K,S})$ in K_S.

12. Applications of GoN: Linear Forms

12.1. Vinogradov's Lemma.

Here is an elementary but ridiculously useful result, apparently first proved by I.M. Vinogradov [Vi27] (and later, presumably independently, by A. Scholz [Sc39]).

Theorem 12.1. *(Vinogradov's Lemma) Let $a, b, n \in \mathbb{Z}^+$ with $n > 1$ and $\gcd(ab, n) = 1$, and let $\alpha \in \mathbb{R}^{>0}$. There are integers x, y, not both zero, such that:*
(i) $ax \equiv by \pmod{n}$,
(ii) $|x| < \alpha,\ |y| \leq \frac{n}{\alpha}$.

Proof. Since $\gcd(a,n) = 1$, there exists $c \in \mathbb{Z}$ with $ac \equiv b \pmod{n}$. Now consider the linear system with defining matrix $C = \begin{bmatrix} c & n \\ 1 & 0 \end{bmatrix}$. Note that $|\det C| = n$. Taking $\epsilon_1 = \alpha$, $\epsilon_2 = \frac{n}{\alpha}$, we have

$$|\det C|(\operatorname{Covol} \mathbb{Z}^2) = n = \epsilon_1 \epsilon_2,$$

and thus by the Linear Forms Theorem, there exists $(X,Y) \in (\mathbb{Z}^2)^\bullet$ such that

$$|L_1(X,Y)| = |cX + nY| \leq \alpha,$$

$$|L_2(X,Y)| = |X| \leq \frac{n}{\alpha}.$$

Put $x = L_1(X,Y) = cX + nY$, $y = L_2(X,Y) = X$, so $(x,y) \in (\mathbb{Z}^2)^\bullet$, $|x| \leq \alpha$, $|y| \leq \frac{n}{\alpha}$ and $ax = a(cX + nY) \equiv bX \equiv by \pmod{n}$. Thus x, y satisfy the desired conclusion except that we currently have $|x| \leq \alpha$ and we want $|x| < \alpha$.
Step 2: If $\alpha \notin \mathbb{Z}$, then $|x| \leq \alpha \iff |x| < \alpha$, and we are done. If $\alpha \in \mathbb{Z}$, then take $0 < \epsilon < \alpha$ and apply the result of Step 1 with $\alpha - \epsilon$ in place of α: there exist integers x and y, not both zero, so that $|x| \leq \alpha - \epsilon < \alpha$ and $|y| \leq \frac{n}{\alpha - \epsilon}$. But for sufficiently small ϵ we have $\lfloor \frac{n}{\alpha - \epsilon} \rfloor = \lfloor \frac{n}{\alpha} \rfloor$ and thus $|y| \leq \lfloor \frac{n}{\alpha} \rfloor \leq \frac{n}{\alpha}$. □

Remark: Actually it is possible to prove Vinogradov's Lemma using less technology than the Convex Body Theorem: in fact, the Pigeonhole Principle suffices!

Proof. It's enough to reprove Step 1 above, since Step 2 is thoroughly elementary. Step 1: Consider $\{S = (i,j) \in \mathbb{Z}^2 \mid 0 \leq i \leq \lfloor \alpha \rfloor,\ 0 \leq j \leq \lfloor \frac{n}{\alpha} \rfloor\}$. Since $\#S = (\lfloor \alpha \rfloor + 1)\left(\lfloor \frac{n}{\alpha} \rfloor + 1\right) > \alpha \cdot \frac{n}{\alpha} = n$, by the Pigeonhole Principle there are distinct elements $(i_1, j_1),\ (i_2, j_2) \in S$ such that

$$ai_1 - bj_1 \equiv ai_2 - bj_2 \pmod{n}.$$

Put $x = i_1 - i_2$, $y = j_1 - j_2$: $(x,y) \neq (0,0)$, $ax \equiv by \pmod{n}$, $|x| \leq \alpha, |y| \leq \frac{n}{\alpha}$. □

Brauer and Reynolds give the following partial generalization of Vinogradov's Lemma to the number field case.

Theorem 12.2. *Let K be an algebraic number field with integer ring $\mathbb{Z}_K$, and let $\mathfrak{m}$ be a nonzero ideal of $\mathbb{Z}_K$ of norm t. Assume condition (BR): for all rational integers $n \in \mathfrak{m}$, $t < n^2$. Then: for $\alpha, \beta \in \mathbb{Z}_K$, the congruence $\alpha x - \beta y \equiv 0 \pmod{m}$ has a solution in* rational integers *x and y, not both in $\mathfrak{m}$, such that $|x|, |y| \leq \sqrt{t}$.*

Proof. By hypothesis, the integers $0, 1, \ldots, \lfloor \sqrt{t} \rfloor$ are pairwise incongruent modulo $\mathfrak{m}$. Letting x and y run through these integers, we get $(\lfloor \sqrt{t} \rfloor + 1)^2 > t$ ordered pairs, hence there exist $(x', y') \neq (x'', y'')$ such that $\alpha x' - \beta y' \equiv \alpha x'' - \beta y'' \pmod{m}$. Put $x = x' - x''$ and $y = y' - y''$. □

Remark: As remarked in [BR51], condition (BR) holds when $\mathfrak{m} = \prod_{i=1}^r \mathfrak{p}_i$ is a product of prime ideals whose norms are distinct rational primes. It also holds when the norm of each $\mathfrak{p}_i$ is either p_i or p_i^2 (for distinct primes p_i) as long as at least one of the norms is prime.

I am not aware of any Diophantine applications of Theorem 12.2.

12.2. **Improvements on Vinogradov: Brauer-Reynolds and Cochrane.**

In a classic 1951 paper, A. Brauer and R.L. Reynolds used pigeonholing arguments to prove a substantial (and useful) generalization of Vinogradov's Lemma.

Theorem 12.3. *(Brauer-Reynolds* [BR51]*) Let $m, n, d \in \mathbb{Z}^+$ with $d > 1$, $m < n$, and let $\lambda_1, \ldots, \lambda_n \in \mathbb{R}^{>0}$ be such that $\lambda_i < d$ for all i and $\lambda_1 \cdots \lambda_n > d^m$. Then for any matrix $A = (a_{ij}) \in M_{m,n}(\mathbb{Z})$, the system $Ax \equiv 0 \pmod d$ has a solution $x = (x_1, \ldots, x_n) \in (\mathbb{Z}^n)^\bullet$ with $|x_i| < \lambda_i$ for all $1 \leq i \leq n$.*

Proof. For $1 \leq i \leq m$, put $y_i = y_i(x_1, \ldots, x_n) = \sum_{j=1}^n a_{ij} x_j$. For $x \in \mathbb{R}$, we denote by x^* the largest integer which is strictly smaller than x. For $1 \leq j \leq n$, letting each x_j take integer values from 0 to λ_j^* gives $\prod_{j=1}^n \lambda_j^* + 1$ sets of n-tuples $(x_1, \ldots, x_n)$. Since $\prod_{j=1}^n \lambda_j^* + 1 \geq \prod_{j=1}^n \lambda_j > d^m$, by the Pigeonhole Principle there are $x' = (x_1', \ldots, x_n') \neq (x_1'', \ldots, x_n'') = x''$ such that for all $1 \leq i \leq m$,

$$y_i' = y_i(x') = a_{i1}x_1' + \ldots + a_{in}x_n' \equiv a_{i1}x_1'' + \ldots + a_{in}x_n'' = y_i(x'') = y_i'' \pmod d.$$

For $1 \leq i \leq n$, put $x_j = x_j' - x_j''$, Then for $1 \leq i \leq m$,

$$y_i(x_1, \ldots, x_n) = a_{i1}(x_1' - x_1'') + \ldots + a_{in}(x_n' - x_n'') \equiv 0 \pmod d.$$

Since for all $1 \leq j \leq n$, x_j', x_j'' both lie in $[0, \lambda_j^*]$, $|x_j| = |x_j' - x_j''| \leq \lambda_j^* < \lambda_j$. □

Corollary 12.4. *(Diagonal Case) Under the hypotheses of the theorem, the system $Ax \equiv 0 \pmod d$ has a solution $x \in (\mathbb{Z}^n)^\bullet$ with $|x_i| \leq d^{\frac{m}{n}}$ for all $1 \leq i \leq n$.*

In 1987, T. Cochrane gives a result which improves upon the corollary.

Theorem 12.5. *(Cochrane* [Co87]*) Let $d, m, n \in \mathbb{Z}^+$ with $m \leq n$, and let $A \in M_{m,n}(\mathbb{Z})$ with rank r and invariant factors $d_1, \ldots, d_r$. There is a nonzero solution to the congruence $Ax \equiv 0 \pmod d$ with*

$$\max |x_i| \leq \frac{d^{\frac{r}{n}}}{\prod_{i=1}^r \gcd(d, d_i)^{\frac{1}{n}}}.$$

12.2.1. *An improvement of Stevens and Kuty.*

Inspired by work of Mordell [Mo51], H. Stevens and L. Kuty [SK68] gave a modest – but useful – improvement of the Brauer-Reynolds Theorem. We present their results (in a slightly different, and very slightly stronger) form here.

Theorem 12.6. *(Linear Pigeonhole Principle) Let G_1 and G_2 be groups (not necessarily commutative, but written additively), and let $\Phi : G_1 \to G_2$ be a homomorphism. Let $S \subset G_1$ be a nonempty subset, and let $D(S) = \{s_1 - s_2 \mid s_1, s_2 \in S\}$ be its difference set. If for a cardinal number κ we have*

$$\#S > \kappa \cdot \#G_2,$$

then there are at least κ nonzero elements of $D(S) \cap \operatorname{Ker} \Phi$.

Proof. If $y \in G_2$ we had $\#(\Phi^{-1}(y) \cap S) \leq \kappa$, then $\#S \leq \kappa \cdot \#G_2$, a contradiction. So there is a subset $S' \subset S$ of cardinality greater than κ such that for all $s_1, s_2 \in S'$, $\Phi(s_1) = \Phi(s_2)$. Fix $s_0 \in S$, and let $S' = S \setminus \{s_0\}$, so $\#S' \geq \kappa$. As s runs through S' the elements $s - s_0$ are distinct in G_1 and such that $\Phi(s - s_0) = \Phi(s) - \Phi(s_0) = 0$. □

For $\alpha \in \mathbb{R}$, we denote by α^* the largest integer which is strictly less than α.

Theorem 12.7. *Let $m, n \in \mathbb{Z}^+$, $d_1, \ldots, d_m \in \mathbb{Z}^\bullet$, $\epsilon_1, \ldots, \epsilon_n \in \mathbb{R}^{>0}$, and suppose*

$$\prod_{j=1}^n \epsilon_j \geq \prod_{i=1}^m |d_i|. \tag{23}$$

Let $A = (a_{ij}) \in M_{m,n}(\mathbb{Z})$ and $j_0 \in \{1, \ldots, n\}$. There is $(x_1, \ldots, x_n) \in (\mathbb{Z}^n)^\bullet$ with
(i) $\sum_{j=1}^n a_{ij} x_j \equiv 0 \pmod{d_i}$ for $1 \leq i \leq m$ and
(ii) $|x_{j_0}| \leq \epsilon_{j_0}$, and $|x_j| < \epsilon_j$ for all $j \neq j_0$.

Proof. Let $G_1 = \mathbb{Z}^n$, $G_2 = \prod_{i=1}^m \mathbb{Z}/d_i\mathbb{Z}$. Let $\alpha : \mathbb{Z}^n \to \mathbb{Z}^m$ be given by $(x_1, \ldots, x_n) \mapsto A(x_1, \ldots, x_n)^t$, let $\beta : \mathbb{Z}^m \to \prod_{i=1}^m \mathbb{Z}/d_i\mathbb{Z}$ be the product of the quotient maps, and let $\Phi = \beta \circ \alpha : G_1 \to G_2$. Let

$$S = \mathbb{Z}^n \cap \left([0, \epsilon_{j_0}) \times \prod_{j \neq j_0} [0, \epsilon_j^*) \right).$$

Then

$$\#S = (\lfloor \epsilon_{j_0} \rfloor + 1) \prod_{j \neq j_0} (\epsilon_j + 1) > \prod_{j=1}^s \epsilon_j \geq \prod_{i=1}^m d_i = 1 \cdot \#G_2,$$

so by Theorem 12.6 there are $s_1 \neq s_2 \in S$ with $\Phi(s_1 - s_2) = 0$. Take $v = s_1 - s_2$. □

Theorem 12.8. *Let $m, n \in \mathbb{Z}^+$, let $d_1, \ldots, d_m \in \mathbb{F}_q[t]^\bullet$, $\epsilon_1, \ldots, \epsilon_n \in \mathbb{N}$, and suppose*

$$\sum_{j=1}^n \epsilon_j \geq \sum_{i=1}^m \deg d_i. \tag{24}$$

Let $A = (a_{ij}) \in M_{m,n}(\mathbb{F}_q[t])$. There is $v = (x_1, \ldots, x_n) \in (\mathbb{F}_q[t]^n)^\bullet$ such that
(i) $\sum_{j=1}^n a_{ij} x_j \equiv 0 \pmod{d_i}$ for $1 \leq i \leq m$ and
(ii) $\deg x_{j_0} \leq \epsilon_{j_0}$, $\deg x_j < \epsilon_j$ for all $j \neq j_0$.

Proof. Let $G_1 = \mathbb{F}_q[t]^n$, $G_2 = \prod_{i=1}^m \mathbb{F}_q[t]/(d_i)$. Let $\alpha : \mathbb{F}_q[t]^n \to \mathbb{F}_q[t]^m$ be given by $(x_1, \ldots, x_n) \mapsto A(x_1, \ldots, x_n)^t$, let $\beta : \mathbb{F}_q[t]^m \to \prod_{i=1}^m \mathbb{F}_q[t]/(d_i)$ be the product of the quotient maps, and let $\Phi = \beta \circ \alpha : G_1 \to G_2$.
Let S be the subset of $\mathbb{F}_q[t]^n$ of n-tuples of polynomials $(x_1, \ldots, x_n)$ such that $\deg x_{j_0} \leq \epsilon_{j_0}$ and $\deg x_j \leq \epsilon_j$ for all $j \neq j_0$. Then

$$\#S = (q^{\epsilon_{j_0}}) \cdot \prod_{j \neq j_0} q^{\epsilon_j + 1} = q \prod_{j=1}^n q^{\epsilon_j} > \prod_{1 \leq j \leq n} q^{\epsilon_j} = 1 \cdot \#G_2,$$

so by Theorem 12.6 there are $s_1 \neq s_2 \in S$ with $\Phi(s_1 - s_2) = 0$. Take $v = s_1 - s_2$. □

12.3. **A Number Field Analogue of Brauer-Reynolds.**

The results of this section are (at most) small variants of those of [Co87].

Lemma 12.9. *Let K be a number field, $N \in \mathbb{Z}^+$, $\Lambda \subset \mathbb{Z}_K^N$ a finite index subgroup. Let $\lambda_1, \ldots, \lambda_N$ be positive real numbers such that $\prod_{i=1}^N \lambda_i > M(K)^N [R^n : \Lambda]$. Then there exists $x = (x_1, \ldots, x_N) \in \Lambda^\bullet$ such that $|N(x_i)| \leq \lambda_i$ for $1 \leq i \leq N$.*

Proof. Recall the canonical embedding $\sigma : K \hookrightarrow \mathbb{R}^n$. Using this we define a canonical embedding $\hat{\sigma} : K^N \hookrightarrow \mathbb{R}^{Nn}$, $(x_1, \ldots, x_N) \mapsto (\sigma(x_1), \ldots, \sigma(x_N))$. Arguing exactly as in the $N = 1$ case we see that $\hat{\sigma}(\mathbb{Z}_K^N)$ is a lattice in $\mathbb{R}^{Nn}$ with covolume

$(2^{-s}|\sqrt{d(K)}|)^N$ and thus $\hat{\sigma}(\Lambda)$ is a lattice with covolume $[\mathbb{Z}_K^N : \Lambda](2^{-s}|\sqrt{d(K)}|)^N$. We define $S_1(\lambda) = S(\lambda_1, \ldots, \lambda_N) \subset \mathbb{R}^{Nn}$ as the set of all $x \in \mathbb{R}^{Nn}$ satisfying

$$|x_{i1}| \cdots |x_{ir}||x_{i(r+1)}^2 + x_{i(r+2)}^2| \cdots |x_{i(n-1)}^2 + x_{in}^2| \leq \lambda_i$$

for $1 \leq i \leq N$. By the AGM inequality, $S_1(\lambda)$ contains the symmetric compact convex body $S_2(\lambda)$ defined by

$$|x_{i1}| + \ldots + |x_{ir}| + 2|x_{i(r+1)}^2 + x_{i(r+2)}^2|^{\frac{1}{2}} + \ldots + 2|x_{i(n-1)}^2 + x_{in}^2|^{\frac{1}{2}} \leq n\lambda_i^{\frac{1}{n}}$$

for all $1 \leq i \leq N$. By Proposition 11.2,

$$\operatorname{Vol} S_2(\Lambda) = \left(2^{r-s}\pi^s \frac{n^n}{n!}\right)^N (\lambda_1 \cdots \lambda_N).$$

Applying Minkowski's Convex Body Theorem, there is a nonzero point of Λ in $S_2(\lambda)$ (hence also in $S_1(\lambda)$ if

$$\operatorname{Vol}(S_2(\lambda)) \geq 2^{Nn} \operatorname{Covol} \hat{\sigma}(\Lambda),$$

i.e., iff

$$(\lambda_1 \cdots \lambda_N) \geq M_K^N [Z_K^N : \Lambda].$$

Since a point in $S_1(\lambda)$ satisfies $|N(x_i)| \leq \lambda_i$ for all i, the result follows. □

Proposition 12.10. *(Cochrane* [Co87]*) Let R be a Dedekind domain. Let $U \in M_{m,n}(R)$ have rank r. For a nonzero ideal $\mathfrak{a}$ of R, set*

$$\Lambda = \{x \in R^n \mid Ux \equiv 0 \pmod{\mathfrak{a}}\}.$$

Then $[R^n : \Lambda] \leq (\#R/\mathfrak{a})^r$.

Proof. Since the rank of a matrix can be defined in terms of vanishing of determinants of minors, when we reduce modulo $\mathfrak{a}$, the rank of U will still be at most r. Since R is a Dedekind domain, $R/\mathfrak{a}$ is a principal ring (not necessarily a domain, but it's okay!) and the linear system can be taken to be in Smith Normal Form over $R/\mathfrak{a}$ (without changing the kernel, up to $R/\mathfrak{a}$-module isomorphism). Thus we have at most r equations of the form $a_i x_i = 0$ in $R/\mathfrak{a}$. Each of these equations corresponds to a subgroup of R^n of index at most $\#(R/\mathfrak{a})$ so the conjunction of them gives a subgroup of index at most $\#(R/\mathfrak{a})^r$. □

Here is a variant of Cochrane's result which is actually easier to prove.

Proposition 12.11. *Let R be a Dedekind domain and $U = (a_{ij}) \in M_{m,n}(R)$. Let $\mathfrak{a}_1, \ldots, \mathfrak{a}_m$ be nonzero ideals of R, and let*

$$\Lambda = \{x \in R^n \mid \forall 1 \leq i \leq m, \sum_{j=1}^n a_{ij}x_j \equiv 0 \pmod{\mathfrak{a}_i}\}.$$

Then

$$[R^n : \Lambda] \leq \prod_{i=1}^m \#(R/\mathfrak{a}_i).$$

Exercise: Prove Proposition 12.11.

Theorem 12.12. *Let K be a number field.*
a) Let $\mathfrak{d}$ be a nonzero ideal of $\mathbb{Z}_K$, and let $U \in M_{m,n}(\mathbb{Z}_K)$ with rank r over K. Let $\lambda_1, \ldots, \lambda_n > 0$ satisfy $\lambda_1 \cdots \lambda_n > M(K)^n |\mathfrak{d}|^r$. There is a nonzero solution to the congruence $Ux \equiv 0 \pmod{\mathfrak{d}}$ with $|N(x_i)| \leq \lambda_i$ for all $1 \leq i \leq n$.
b) Let $\mathfrak{a}_1, \ldots, \mathfrak{a}_m$ be nonzero ideals of $\mathbb{Z}_K$. Let $\lambda_1, \ldots, \lambda_n > 0$ such that

$$\prod_{i=1}^{n} \lambda_i > M(K)^n \prod_{i=1}^{m} |\mathfrak{a}_i|.$$

Then there is a nonzero $x \in \mathbb{Z}_K^n$ such that

$$\sum_{j=1}^{n} a_{ij} x_j \equiv 0 \pmod{\mathfrak{a}_j} \ \forall 1 \leq i \leq m$$

and

$$|N(x_i)| \leq \lambda_i \ \forall 1 \leq i \leq n.$$

Proof. a) Combine Proposition 12.10 and Lemma 12.9.
b) Combine Proposition 12.11 and Lemma 12.9. □

13. Applications of GoN: Diophantine Approximation

13.1. **Around Dirichlet's Theorem.**

Minkowski's Theorem on Linear Forms is closely related to **Diophantine Approximation**. The following is the most basic result in this area.

Theorem 13.1. *(Dirichlet) Let $\alpha \in \mathbb{R}$ and $Q \in \mathbb{Z}^+$. There are $p, q \in \mathbb{Z}$ with $1 \leq p \leq Q$ such that $|\alpha - \frac{p}{q}| \leq \frac{1}{q(Q+1)}$.*

Proof. Consider the pair of linear forms

$$L_1(x_1, x_2) = x_1 - \alpha x_2.$$

$$L_2(x_1, x_2) = x_2.$$

The corresponding matrix C has determinant 1. Let $\Lambda = \mathbb{Z}^2$. For any $\epsilon_1, \epsilon_2 > 0$ such that $\epsilon_1 \epsilon_2 \geq 1$, by Theorem 9.3 there are $p, q \in \mathbb{Z}$, not both zero, such that

$$|p - \alpha q| \leq \epsilon_1, \ |q| \leq \epsilon_2.$$

Note that for $\epsilon_1 < 1$ the above inequalities imply $q \neq 0$. Fix $\theta \in (0,1)$ and take $\epsilon_1 = \frac{1}{Q+\theta}$, $\epsilon_2 = Q + \theta$. Then $|p - \alpha q| \leq \frac{1}{Q+\theta}$ and $|q| \leq \lfloor Q + \theta \rfloor = Q$. But there are only finitely many $(p, q) \in \mathbb{Z}^2 \setminus \{0\}$ satisfying $|q| \leq Q + 1$, $|p - \alpha q| \leq \frac{1}{Q}$, so if the above inequalities hold for all $\theta < 1$, there must exist $(p, q) \in \mathbb{Z}^2$ such that $|p - \alpha q| \leq \frac{1}{Q+1}$, $|q| \leq Q$. These p and q satisfy the conclusion of the theorem. □

Corollary 13.2. *If $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, then for infinitely many nonzero integers y, there exists an integer x such that*

$$\left| \alpha - \frac{x}{y} \right| < \frac{1}{y^2}. \tag{25}$$

Exercise: a) Prove Corollary 13.2.
b) Show that conversely, if $\alpha \in \mathbb{Q}$, there are only finitely many nonzero integers y for which there exists an integer x such that (25) holds.

It is well known that Dirichlet's Theorem can be proven by a simple Pigeonhole Principle argument (indeed, this seems to be the first use of the Pigeonhole Principle to prove a nontrivial result, and in some circles one speaks of the *Dirichlet Box Principle* instead of the Pigeonhole Principle). However, almost the same argument can be used to prove the following result on simultaneous approximation.

For any real number α, we denote by $||\alpha||$ the distance to the nearest integer.

Theorem 13.3. *Let $M, N \in \mathbb{Z}^+$, $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ and let $L_1(x), \ldots, L_M(x)$ be linear forms: $L_m(x) = \sum_{n=1}^N a_{mn} x_n$. Let $Q \in \mathbb{Z}^+$. Then there exists $v = (v_1, \ldots, v_N) \in (\mathbb{Z}^N)^\bullet$ such that $\max_n |v_n| \leq Q$ and $\max_{1 \leq m \leq M} ||L_m(v)|| \leq \frac{1}{(Q+1)^{N/M}}$.*

Proof. (Burger [Bu]) Let $A = (a_{mn}) \in M_{MN}(\mathbb{R})$. Define $C \in \mathrm{GL}_{M+N}(\mathbb{R})$ by

$$C = \begin{bmatrix} 1_M & -A \\ 0 & 1_N \end{bmatrix},$$

so that $\det C = 1$. Now we apply Minkowski's Linear Forms Theorem in a context directly generalizing the one in the proof of Theorem 13.1: fix $\theta \in (0,1)$, take $\Lambda = \mathbb{Z}^{M+N}$ and

$$\epsilon_1 = \ldots = \epsilon_M = \frac{1}{(Q+\theta)^{\frac{N}{M}}}, \ \epsilon_{M+1} = \ldots = \epsilon_N = Q + \theta.$$

The remaining details are left to the reader. □

For later use, we record separately the result obtained by setting $N = 1$.

Corollary 13.4. *Let $M, n \in \mathbb{Z}^+$, $M > 1$, $\theta_1, \ldots, \theta_n \in \mathbb{R}$. Then there are integers $\ell_1, \ldots, \ell_n, m$ with $0 < m < M$ and $|m\theta_j - \ell_j| \leq \frac{1}{M^{\frac{1}{n}}}$ for all $1 \leq j \leq n$.*

13.2. **The Best Possible One Variable Approximation Result.**

We begin with the following result, a version of Theorem 9.15 with an additional linear condition.

Theorem 13.5. *(Siegel) Let $q(x,y) = L_1(x,y)L_2(x,y)$ be an indefinite real binary quadratic form, of Discriminant $\Delta > 0$. We suppose that q is* not *H-equivalent to $q_2 = x^2 - xy - y^2$. Then for any $\epsilon > 0$, there is $(x,y) \in (\mathbb{Z}^2)^\bullet$ such that*

$$|q(x,y)| < \sqrt{\frac{\Delta}{5}}$$

and

$$|L_1(x,y)| < \epsilon.$$

Proof. ... □

Theorem 13.6. *Let $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Then for infinitely many nonzero integers y, there exists an integer x such that*

$$\left| \alpha - \frac{x}{y} \right| < \frac{1}{\sqrt{5} y^2}.$$

Proof. (Siegel) Let $L_1(x,y) = (x-\alpha y)$, $L_2(x,y) = y$, and put $q(x,y) = L_1(x,y)L_2(x,y) = xy - \alpha y^2$. Then q is indefinite with Discriminant 1. Further, we claim that q is not integrally equivalent to any scalar multiple of q_2. Indeed, any scalar multiple of q_2 has the following property: for all $(x_1,y_1),(x_2,y_2) \in \mathbb{Z}^2$ such that $q(x_2,y_2) \neq 0$, $\frac{q(x_1,y_1)}{q(x_2,y_2)} \in \mathbb{Q}$, but – since α is irrational – taking $(x_1,y_1) = (1,0)$ and $(x_2,y_2) = (0,1)$ shows that q does not have this property. Thus Theorem 13.5 applies: for any $\epsilon > 0$, there are integers x, y, not both zero, such that

$$|q(x,y) = |y(x-\alpha y)| < \frac{1}{\sqrt{5}}$$

and $|x - \alpha y| < \epsilon$. Taking $\epsilon < 1$ forces $y \neq 0$, so we may divide through by y to get

$$\left|\frac{x}{y} - \alpha\right| < \frac{1}{\sqrt{5}y^2}.$$

Since we may take ϵ as small as we like, there are infinitely many choices of y. □

Remark: Theorem 13.6 appears in many texts, but it is usually proved using continued fractions, e.g. [HW6ed, Thm. 193]. We are indebted to Siegel for allowing us to maintain our "no continued fractions policy" so far inside enemy terrain.

The following simple result shows that Theorem 13.6 is *sharp* in the sense that there are some irrational numbers α for which the constant $\frac{1}{\sqrt{5}}$ cannot be improved.

Proposition 13.7. *Let $\alpha = \frac{1-\sqrt{5}}{2}$. For any $A > \sqrt{5}$, the inequality*

$$\left|\alpha - \frac{x}{y}\right| < \frac{1}{Ay^2}$$

has only finitely many solutions.

Proof. (Hardy-Wright) Suppose not. Then there are infinitely many pairs (x,y) with $\alpha = \frac{x}{y} + \frac{\delta}{y^2}$ with $|\delta| < \frac{1}{A} < \frac{1}{\sqrt{5}}$. Then

$$\frac{\delta}{y} = y\alpha - x, \ \frac{\delta}{y} - \frac{\sqrt{5}y}{2} = \frac{y}{2} - x,$$

so

$$\frac{\delta^2}{y^2} - \sqrt{5}\delta = \left(\frac{y}{2} + x\right)^2 - \frac{5y^2}{4} = x^2 + xy - y^2. \tag{26}$$

For sufficiently large y, the left hand side of (26) is less than one in absolute value, whereas the right hand side is always an integer, hence $x^2 + xy - y^2 = 0$. But this is impossible because the Discriminant of $q(x,y) = x^2 + xy - y^2$ is $\sqrt{5}$, which is irrational, so $q(x,y) = 0$ has as its only integral solution $x = y = 0$. □

13.3. **The Markoff Chain.**

To do: State the main result of [Ca49].

14. Applications of GoN: Euclidean Rings

Let R be a domain with fraction field K. Let $|\cdot| : R \to \mathbb{N}$ be a **multiplicative norm function**: $\forall x, y \in R$,

(MN1) $|x| = 0 \iff x = 0$;
(MN2) $|x| = 1 \iff x \in R^\times$;
(MN3) $|xy| = |x||y|$.

Exercise: Show that a multiplicative norm extends uniquely to a map $|\cdot| : K \to \mathbb{Q}^{\geq 0}$ such that $|\frac{x}{y}| = \frac{|x|}{|y|}$.

A normed domain $(R, |\cdot|)$ is **Euclidean** if for all $a, b \in R$ with $b \neq 0$, there are $q, r \in R$ with $a = qb + r$ and $|r| < |b|$.

Exercise: Show that the Euclidean condition is equivalent to: for all $x \in K$ there is $y \in R$ with $|x - y| < 1$.

Exercise: We say two norms $|\cdot|_1$, $|\cdot|_2$ on a domain R are **equivalent** if there is some $\alpha \in \mathbb{R}^{>0}$ such that $|\cdot|_2 = |\cdot|_1^\alpha$. Show that this is an equivalence relation and that being Euclidean depends only on the equivalence class.

There are two classical cases:

Exercise: a) Let $R = \mathbb{Z}$ endowed with the standard norm (i.e., the usual absolute value coming from $\mathbb{R}$). Show that $(R, |\cdot|)$ is Euclidean.
b) Let $\mathbb{F}_q$ be a finite field, and let $R = \mathbb{F}_q[t]$. For $x \in \mathbb{F}_q[t]^\bullet$, put $|x| = q^{\deg x}$. Show that $(R, |\cdot|)$ is Euclidean.

For $y \in K$, we set

$$E(R, y) = \inf_{x \in R} |x - y|,$$
$$E(R) = \sup_{y \in K} E(R, y).$$

Now we make some basic observations:

- $E(R, y)$ depends only on the class of y in K/R.
- R is Euclidean iff $E(R, y) < 1$ for all $y \in K/R$.
- R is Euclidean if $E(R) < 1$.
- R is not Euclidean if $E(R) > 1$.
- If $E(R) = 1$, R is Euclidean iff the supremum in the definition of $E(R)$ is *not* attained.

We also put

$$E_1(R) = E(R),$$
$$C_1(R) = \{y \in K/R \mid E(R, y) = E_1(R)\}$$

and

$$E_2(R) = \sup\{E(R, y) \mid y \notin C_1(R)\},$$
$$C_2(R) = \{y \in K/R \mid E(R, y) = E_2(R),$$

$$E_3(R) = \sup\{E(r,y) \mid y \notin C_2(R),$$

and so forth: we get a sequence

$$E(R) \geq E_2(R) \geq E_3(R) \geq \dots.$$

If $E_2(R) < E_1(R)$ we say $E_1(R)$ is **isolated**.

In this section we will restrict to an important classical case (though the author is equally interested in other cases...) $R = \mathbb{Z}_K$, the ring of integers of a number field K. As usual, let $n = [K:\mathbb{Q}] = r + 2s$, where r is the number of real embeddings and s is the number of complex embeddings of K. Also put $r_u = r + s - 1$, the rank of the unit group $\mathbb{Z}_K^\times$. We take as a norm function $|x| = |N_{K/\mathbb{Q}}(x)|$, where the norm on the right hand side is the usual absolute value on $\mathbb{Q}$. In this case, whether R is Euclidean is a GoN problem that can be studied via the embedding of $\sigma : K \to \mathbb{R}^n \cong \mathbb{R}^r \times \mathbb{C}^s$ in Euclidean space encountered in § 11. For $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ we put

$$|x| = |\prod_{i=1}^n x_i|.$$

Thus if we choose a $\mathbb{Q}$-basis $\alpha_1, \dots, \alpha_n$ of K, we have

$$|(x_1, \dots, x_n) = \prod_{i=1}^n \sum_{j=1}^n x_j \sigma_i(\alpha_j).$$

Elements of $\mathbb{R}^n$ in the image of $\sigma(K)$ are called **rational points**; the others are **irrational**.

This leads to a "geometric approach" to showing that $\mathbb{Z}_K$ is Euclidean. We define, for all $y \in \mathbb{R}^n$,

$$E(R,y) = \inf_{x \in R} |x - y|$$

and

$$\overline{E}(R) = \sup_{y \in \mathbb{R}^n} E(R,y).$$

Then clearly

$$E(R) \leq \overline{E}(R),$$

so if $\overline{E}(R) < 1$, R is Euclidean. Here is an easy instance of this.

Proposition 14.1. *Let D be a squarefree negative integer which is not congruent to 1 modulo 4, and put $K = \mathbb{Q}(\sqrt{D})$, $R = \mathbb{Z}_K = \mathbb{Z}[\sqrt{D}]$.*
a) We have $E(R) = \overline{E}(R) = \frac{1+|D|}{4}$.
b) Thus R is Euclidean iff $D \in \{-1, -2\}$.

Theorem 14.2. *a) (Barnes-Swinnerton-Dyer, Cerri) For all number fields K, $E(R) = \overline{E}(R)$.*
b) (Cerri) If $n \geq 3$ and $r_u \geq 2$, then there is $y \in K$ such that

$$E(R) = \overline{R}(R) = E(R,y).$$

In particular, in this case R is Euclidean $\iff$ $E(R) < 1$.

15. Applications of GoN: Representation Theorems for Quadratic Forms

15.1. **Reminders on integral quadratic forms.**

An **integral quadratic form** is a homogeneous polynomial of degree 2 with $\mathbb{Z}$-coefficients in N variables. Such a form may be written as $\sum_{1 \leq i \leq j \leq N} a_{ij} x_i x_j$ with $a_{ij} \in \mathbb{Z}$. We may also define a symmetric matrix M such that for $x = (x_1, \ldots, x_N)$,

$$q(x) = x^T M x.$$

There is a small twist here: in order for the bookkeeping to work out correctly, we must take $M(i,j)$ to be a_{ij} if $i = j$ but $\frac{a_{ij}}{2}$ if $i \neq j$. Thus the matrix M has integer entries on the main diagonal, but off of the main diagonal the entries need only be half-integers. The simplest example of this is the binary quadratic form

$$q(x,y) = x^2 + xy + y^2, \tag{27}$$

with defining matrix

$$\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}.$$

We will restrict to **nondegenerate** quadratic forms here, i.e., ones for which the determinant of the defining matrix is nonzero.

When the matrix M has $\mathbb{Z}$-entries one says q is **classically integral** or an **integer matrix** form, as opposed to being **integral** or **integer valued**.

A diagonal integral quadratic form – i.e., one with $a_{ij} = 0$ for all $i \neq j$ – is necessarily classically integral. For this and other reasons diagonal forms are especially nice to work with, so our first examples (and theorems!) on quadratic forms concern diagonal ones, although we certainly aspire to move beyond this case eventually. Note also that we consider integral quadratic forms up to **equivalence**: we say q_1 and q_2 are integrally equivalent if we can get from one quadratic form to another by a change of variables $A \in \mathrm{GL}_N(\mathbb{Z})$. In terms of matrices this comes out not as similarity but **congruence**:[10]

$$M_{q_2} = A^T M_{q_1} A.$$

Among other things, this equivalence relation raises the prospect of replacing a non-diagonal form with an equivalent diagonal form, or in a word, **diagonalizing**. Recall the following elementary but important fact: any quadratic form with coefficients in a *field* of characteristic different from 2 may be diagonalized. In particular any integral quadratic form q may be diagonalized over $\mathbb{Q}$ or over $\mathbb{R}$. But in fact "most" integral quadratic forms cannot be diagonalized over $\mathbb{Z}$.

Exercise:
a) Let $f(x_1, \ldots, x_n), g(x_1, ., \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ be two integral quadratic forms. Show: if $f \sim g$ and f is classically integral, then g is classically integral.
b) Show that $f(x,y) = x^2 + xy + y^2$ cannot be diagonalized over $\mathbb{Z}$.

[10]Here **congruence** is just a name, somewhat old-fashioned at that. Other than that congruence – like similarity! – of matrices is an equivalence relation, it has nothing much to do with other notions of congruence studied in algebra.

Let $q = q(x_1, \ldots, x_n) \in \mathbb{R}[x_1, \ldots, x_n]$ be a real quadratic form. We say q is **positive definite** if for all $x \in \mathbb{R}^n$, we have $q(x) \geq 0$ and $q(x) = 0 \iff x = 0$. (We say q is **negative definite** if for all $x \in \mathbb{R}^n$, we have $q(x) \leq 0$ and $q(x) = 0 \iff x = 0$. But q is negative definite iff $-q$ is positive definite, so negative definite forms need not be studied for their own sake.) We say q is **indefinite** if there are $x, y \in \mathbb{R}^n$ with $q(x) > 0$ and $q(y) < 0$.

Exercise: A real quadratic form $q(x_1, \ldots, x_n)$ is **positive semidefinite** if $q(x) \geq 0$ for all $x \in \mathbb{R}^n$. Show that a nondegenerate positive semidefinite quadratic form is positive definite.

Exercise: Let $q \in \mathbb{R}[x_1, \ldots, x_n]$ be an anisotropic real quadratic form such that $q(x) \geq 0$ for all $x \in \mathbb{Z}^n$. Show that q is positive definite.

From the geometric perspective, positive and negative definite quadratic forms are very different.

Exercise: a) Show that a real quadratic form is positive definite iff it is $\mathbb{R}$-equivalent to a diagonal form $a_1 x_1^2 + \ldots + a_n x_n^2$ with $a_1, \ldots, a_n > 0$.
b) Suppose you are given a non-diagonal integral quadratic form q. Part a) gives a procedure for checking whether q is positive definite: diagonalize it, and see whether the diagonal coefficients are all positive. Is this actually the fastest procedure in practice?[11]

Exercise: Let q be a positive definite real quadratic form. Show that $q : \mathbb{R}^N \to \mathbb{R}$ is a symmetric, convex distance function whose level sets are ellipsoids.

Exercise: Let q be an indefinite real quadratic form. Show that $q : \mathbb{R}^N \to \mathbb{R}$ is a symmetric, non-convex pseudo-distance function with non-compact level sets.

An integral quadratic form q is **sign universal** if:
(i) $q(\mathbb{R}^N) = \mathbb{R}^{\geq 0}$ and $q(\mathbb{Z}^N) = \mathbb{Z}^{\geq 0}$, or
(ii) $q(\mathbb{R}^N) = \mathbb{R}$ and $q(\mathbb{Z}^N) = \mathbb{Z}$, or
(iii) $q(\mathbb{R}^N) = \mathbb{R}^{\leq 0}$ and $q(\mathbb{Z}^N) = \mathbb{Z}^{\leq 0}$.

In particular, a positive definite integral quadratic form is sign universal iff it represents all non-negative integers, which is the largest subset of the integers it could conceivably represent.

Here are three important facts about sign universal positive definite forms.

Theorem 15.1. *Let q be a positive definite integral quadratic form in N variables.*
a) If $N \leq 3$ then q is not *sign universal.*
b) (Conway-Schneebeger [Con00]*, Bhargava* [Bh00]*) If q is classically integral, it is sign universal iff it $\mathbb{Z}$-represents all integers from 1 to 15.*

[11] There is also a criterion due to Sylvester involving positivity of principal minors. I vaguely suspect this may be faster than diagonalization, but I have never given it serious thought.

c) (Bhargava-Hanke [BH]*) If q is integral, it is sign universal iff it $\mathbb{Z}$-represents all integers from 1 to 290.*

Proof. a) We will show a stronger result: no positive definite ternary quadratic form $q(x,y,z)_{/\mathbb{Q}}$ can $\mathbb{Q}$-represent every positive integer. Note that if such a form $\mathbb{Q}$-represents all positive integers, then for all $a, b \in \mathbb{Z}^+$ it $\mathbb{Q}$-represents ab and thus also $(b^{-1})^2 ab = \frac{a}{b}$, i.e., it $\mathbb{Q}$-represents every positive rational number.

We may diagonalize q over $\mathbb{Q}$. Moreover, we may replace q by $(\operatorname{disc} q)q$, giving a positive universal definite form of discriminant (squareclass) 1. Thus

$$q = -ax^2 - by^2 + abz^2$$

for some $a, b \in \mathbb{Q}^\times$: that is, $q = n_0$ is the **ternary norm form** of the quaternion algebra $\langle a, b\rangle$ in the sense of [NCA, §5.5]. Now q is positive definite, so it is anisotropic over $\mathbb{Q}$. By the Hasse-Minkowski theory it is also anisotropic over $\mathbb{Q}_p$ for at least one prime p. We claim that in fact q $\mathbb{Q}_p$-represents every nonzero element of $\mathbb{Q}_p$. For this, it suffices to show that every square class in $\mathbb{Q}_p$ contains a positive rational number, which for instance holds by weak approximation. Now we use the following fundamental fact [NCA, Thm. 94]: if $q(x,y,z)$ is an anisotropic ternary quadratic form with discriminant 1 over *any* field K of characteristic different from 2, then the quaternary quadratic form $q'(x,y,z,w) = q(x,y,z) + w^2$ is also anisotropic. Applying this in our case we get a contradiction: since q is universal over $\mathbb{Q}_p$, there exist $(x,y,z) \in \mathbb{Q}_p^3$ such that $q(x,y,z) = -1$, and then $q'(x,y,z,1) = 0$, so q' is isotropic over $\mathbb{Q}_p$.
b) This is the **Conway-Schneeberger 15 Theorem**. See [Bh00] for a proof.
c) This is the **Bhargava-Hanke 290 Theorem**. □

Remark 7. *Part a) is a classical result. I don't know how far back it goes, but papers in (e.g.) the 1930's mention the result without any citation or proof.*

It is also interesting to consider universal ternary forms over rings of integers of number fields $K \supsetneq \mathbb{Q}$. For instance suppose $K = \mathbb{Q}(\sqrt{d})$ with $d > 0$ and $q(x,y,z)$ is positive definite at both the real places of K. Then the above argument does not quite succeed in showing that q cannot K-rationally represent all totally positive elements of K: the conic $q(x,y,z)$ may be anisotropic only at the two real places of K and at no finite place. And in fact there are known examples of K and ternary quadratic forms $q_{/\mathbb{Z}_K}$ which $\mathbb{Z}_K$-represent every totally positive element of $\mathbb{Z}_K$! This is an active research area.

Part b) was proven in a graduate course taught by J.H. Conway at Princeton in 1993 (see [Con00]*) but only written up by M. Bhargava about ten years later. The original proof required some nontrivial computer calculations, but Bhargava's proof* [Bh00] *is beautifully conceptual.*

The result of part c) was announced by M. Bhargava and J.P. Hanke in 2005.

Exercise: Let $n \in \mathbb{Z}^+$ and consider the form

$$q_{4,n} = x^2 + y^2 + nz^2 + nw^2.$$

a) Show that $q_{4,n}$ $\mathbb{Z}$-represents 3 iff $1 \leq n \leq 3$.
b) Use Theorem 15.1 to show that for $1 \leq n \leq 3$, the form $q_{4,n}$ is sign universal.

Exercise: Let $1 \leq a \leq b \leq c \leq d \in \mathbb{Z}^+$, and consider the quadratic form

$$q = ax^2 + by^2 + cz^2 + dw^2.$$

Suppose that q is sign universal.
a) Explain why Theorem 15.1a) shows, without any explicit calculation, that the number of tuples (a, b, c, d) such that q is sign universal must be finite.
b) (Ramanujan [Ra17]) Complete the following steps to obtain an explicit upper bound on the set of (a, b, c, d) such that q is sign universal.
(i) Show $a = 1$.
(ii) Show $b \leq 2$.
(iii) If $(a, b) = (1, 1)$, show $c \leq 2$.
(iv) If $(a, b, c) = (1, 1, 1)$, show $d \leq 7$.
(v) If $(a, b, c) = (1, 1, 2)$, show $d \leq 14$.
(vi) If $(a, b) = (1, 2)$, show $c \leq 5$.
(vii) If $(a, b, c) = (1, 2, 2)$, show $d \leq 7$.
(viii) If $(a, b, c) = (1, 2, 3)$, show $d \leq 10$.
(ix) If $(a, b, c) = (1, 2, 4)$, show $d \leq 14$.
(x) If $(a, b, c) = (1, 2, 5)$, show $d \leq 10$.
c) For each of the finitely many forms permitted by part b), apply Theorem 15.1b) to determine whether or not they are universal. You should arrive at the following **Ramanujan-Dickson** list of 54 forms:

$$[1,1,1,1],[1,1,1,2],[1,1,1,3],[1,1,1,4],[11,1,5],[1,1,1,6],[1,1,1,7],[1,1,2,13],$$

$$[1,1,2,2],[1,1,2,3],[1,1,2,4],[1,1,2,5],[1,1,2,6],[1,1,2,7],[1,1,2,8],$$

$$[1,1,2,9],[1,1,2,10],[1,1,2,11],[1,1,2,12],[1,1,2,14],[1,1,3,3],[1,1,3,4],[1,1,3,5],$$

$$[1,1,3,6],[1,2,2,2],[1,2,2,3],[1,2,2,4],[1,2,2,5],[1,2,2,6],[1,2,2,7],[1,2,3,3],$$

$$[1,2,3,4],[1,2,3,5],[1,2,3,6],[1,2,3,7],[1,2,3,8],[1,2,3,9],[1,2,3,10],[1,2,4,4],$$

$$[1,2,4,5],[1,2,4,6],[1,2,4,7],[1,2,4,8],[1,2,4,9],[1,2,4,10],[1,2,4,11],[1,2,4,12],$$

$$[1,2,4,13],[1,2,4,14],[1,2,5,10],[1,2,5,6],[1,2,5,7],[1,2,5,8],[1,2,5,9].$$

Remark: The above list is Dickson's list [Dic27]. Ramanujan's list [Ra17] included the above 54 forms together with the quadratic form $[1, 2, 5, 5]$. However this form does not represent 15! Interestingly, neither Ramanujan nor Dickson characterized their results in terms of representation up to 15. However, this characterization does appear in a paper of P. Halmos [Ha38].[12]

The 15 Theorem leads to a complete enumeration of sign universal positive definite classically integral quaternary forms: there are precisely 204 such forms (up to $\mathbb{Z}$-equivalence). Such a classification was done in the 1948 thesis of M. Willerding but was never published. In fact Willerding's thesis seems not to have been carefully read: although the methods employed are in principle correct, the final tabulation is off by quite a lot: as described in [Bh00] she tallied 178 such forms, and in fact missed 36 universal forms, listed one universal form twice, and listed 9 non-universal forms. Similarly, the 290 Theorem leads to an enumeration of all sign universal positive definite integral quaternary forms: there are 6436 such forms.

[12]Halmos's *first* paper. I learned about by reading his *Automathography*: highly recommended!

15.2. **An application of Hermite's Bound.**

Following Gerstein [G, §7.6] we give a useful application of the Hermite constant to quadratic form theory.

Theorem 15.2. *Let $1 \leq n \leq 7$, and let $q(x_1, \ldots, x_n)$ be an anisotropic classically integral quadratic form with $\det q \in \{\pm 1\}$. Then q is integrally equivalent to either $x_1^2 + \ldots + x_n^2$ or $-x_1^2 - \ldots - x_n^2$.*

Proof. By Theorem 10.2, since $n \leq 7$ we have $\gamma_n < 2$. (We note in passing that Hermite's own estimate gives this for $n \leq 5$. That it holds for $n = 6$ and $n = 7$ are classical, but deep, results of Blichfeldt.) It follows that q represents ± 1, so

$$q \sim a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n + \ldots + a_{22}x_2^2 + \ldots + a_{nn}x_n^2,$$

with $|a_{11}| = 1$. The change of variables $y_j = x_j - \frac{a_{1j}}{a_{11}}$ – note first that this is really just completing the square and second that the fact that f is classically integral is being used here – gives

$$f \sim a_{11}x_1 \oplus f_2(x_2, ..., x_n),$$

and f_2 is anisotropic with $\det f_2 \in \{\pm 1\}$. By induction we get

$$f \sim \epsilon_1 x_1^2 + \ldots + \epsilon_n x_n^2$$

with $\epsilon_1, \ldots, \epsilon_n \in \{\pm 1\}$. Since f is anisotropic, all of the signs must be the same. □

Remark 8. *a) The hypothesis that f is anisotropic is needed here: e.g. $f(x, y) = 2xy$ is classically integral, has determinant -1 and is clearly not of the form $ax^2 + by^2$: we would have to have $|a| = |b| = 1$ and then f would represent ± 1.*
*b) Since $\gamma_8 = 2$, the proof does not go through for $n \geq 8$. Indeed, starting in dimension 8 there are quadratic forms q which are positive definite, have $\operatorname{disc} q = 1$ and have $q(\mathbb{Z}^n) \subset 2\mathbb{Z}$, namely we have the E_8-**lattice**.*

15.3. **The Two Squares Theorem.**

Here we present a geometry of numbers proof of the Two Squares Theorem with the minimal possible number theoretic background (which is quite minimal indeed).

First, $2 = 1^2 + 1^2$. So we may, and shall, restrict our attention to primes $p > 2$.

Lemma 15.3. *(Quadratic character of -1)*
Let $p > 2$ be a prime number. The following are equivalent:
(i) There exists $u \in \mathbb{Z}$ such that $u^2 \equiv -1 \pmod{p}$.
(ii) $p \equiv 1 \pmod{4}$.

Proof. Although we could get away with even less, let us make use the following standard fact from undergraduate algebra: the group of units $U(p) = (\mathbb{Z}/p\mathbb{Z})^\times$ of the finite field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ is cyclic [A2.5, Cor. 10]. Moreover $\#U(p) = \#\mathbb{F}_p^\bullet = p - 1$ is even, so there is a unique element of order 2, namely -1.

An element $u \in U(p) = \mathbb{F}_p^\times$ with $u^2 = -1$ has order 4 in the group $U(p)$, and conversely for any element u of order 4, u^2 has order 2 so $u^2 = -1$. Thus the existence of an element squaring to -1 is equivalent to the existence of an element of order 4 in the cyclic group of order $p-1$, and for this it is necessary and sufficient that $p - 1$ be divisible by 4, i.e., $p \equiv 1 \pmod{4}$. □

Lemma 15.4. *No prime number $p \equiv 3 \pmod 4$ is a sum of two integer squares.*

Proof. Indeed, suppose $p = x^2 + y^2$ with $x, y \in \mathbb{Z}$. Reducing modulo 4 we get $3 \equiv x^2 + y^2 \pmod 4$. Since $0^2 \equiv 2^2 \equiv 0 \pmod 4$ and $1^2 \equiv 3^2 \equiv 1 \pmod 4$, the squares modulo 4 are 0 and 1, and thus the sums of two squares are $0 + 0 = 0$, $0 + 1 = 1$, $1 + 1 = 2$. So 3 is not a sum of two squares in $\mathbb{Z}/4\mathbb{Z}$, contradiction. □

Theorem 15.5. *(Two Squares Theorem for Primes)*
A prime number p is a sum of two integer squares iff $p = 2$ or $p \equiv 1 \pmod 4$.

Proof. It remains to show: a prime $p \equiv 1 \pmod 4$ is a sum of two squares. For such a p, by Lemma 15.3 there is $u \in \mathbb{Z}$ with $u^2 \equiv -1 \pmod p$. Let

$$M := \begin{bmatrix} p & u \\ 0 & 1 \end{bmatrix}.$$

We have $\det(M) = p$, so $\Lambda := M\mathbb{Z}^2$ defines a lattice in $\mathbb{R}^2$ with

$$\operatorname{Covol}(\Lambda) = \det(M) \operatorname{Covol}(\mathbb{Z}^2) = p.$$

If $(t_1, t_2) \in \mathbb{Z}^2$ and $(x_1, x_2)^t = M(t_1, t_2)^t$, then

$$x_1^2 + x_2^2 = (t_1 p + t_2 u)^2 + t_2^2 \equiv (u^2 + 1)t_2^2 \equiv 0 \pmod p.$$

Now let

$$\Omega = B_0(\sqrt{2p}) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 2p\}$$

be the open ball of radius $2\sqrt{p}$ about the origin in $\mathbb{R}^2$. We have

$$\operatorname{Vol} \Omega = \pi(\sqrt{2p})^2 = 2\pi p > 4p = 2^2 \operatorname{Covol} \Lambda,$$

so by Minkowski's Theorem Mark II there exists $(x_1, x_2) \in \Lambda$ with

$$0 < x_1^2 + x_2^2 < 2p.$$

Since $p \mid x_1^2 + x_2^2$, the only possible conclusion is

$$x_1^2 + x_2^2 = p.$$

□

Lemma 15.6. *(Brahmagupta-Fibonacci Identity) For any integers a, b, c, d we have*

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2.$$

Proof. An immediate application of Littlewood's Principle: all algebraic identities are trivial to prove (though not necessarily to discover).[13] □

Lemma 15.7. *a) For any field F, TFAE:*
(i) There exist $x, y \in F$, not both zero, such that $x^2 + y^2 = 0$.
(ii) There exists $i \in F$ with $i^2 = -1$.
b) For a prime number p, the equivalent conditions of part a) hold for the field $\mathbb{F}_p$ iff $p = 2$ or $p \equiv 1 \pmod 4$.

Proof. a) (i) $\implies$ (ii): Without loss of generality $y \neq 0$, and then $-1 = \left(\frac{x}{y}\right)^2$.
(ii) $\implies$ (i): If $i^2 = -1$ then $i^2 + 1 = 0$.
b) This is immediate from the proof of Lemma 15.3. □

[13] Of course the sufficiently learned reader will know more insightful proofs, e.g. using the multiplicativity of the norm function on the ring of Gaussian integers.

Lemma 15.8. *Let $p \equiv 3 \pmod 4$ be a prime number. Then for any integers x and y – not both zero – $\mathrm{ord}_p(x^2+y^2)$ is even: i.e., the largest power of p which divides x^2+y^2 is even.*

Proof. Since zero is even, we may assume that $p \mid x^2+y^2$. Reducing mod p and applying Lemma 15.7 we deduce that there are $X, Y \in \mathbb{Z}$ such that $x = pX$, $y = pY$, so $x^2+y^2 = p^2(X^2+Y^2)$. An evident induction argument finishes the proof. □

Combinig the above results, one deduces the following theorem.

Theorem 15.9. *(Full Two Squares Theorem) For a positive integer n, TFAE:*
(i) n is a sum of two integer squares.
(ii) For all primes $p \equiv 3 \pmod 4$, $\mathrm{ord}_p(n)$ is even.

15.4. **Binary Quadratic Forms.**

An integral binary quadratic form is a polynomial

$$q(x,y) = Ax^2 + Bxy + Cy^2$$

with $A, B, C \in \mathbb{Z}$. This form corresponds to a matrix

$$\begin{bmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{bmatrix}$$

with determinant $AC - \frac{B^2}{4}$. Elsewhere in these notes we call this determinant the *discriminant* of the quadratic form q. But here we run afoul of the variation in terminology and notation in this subject: for binary quadratic forms, it is traditional to call the quantity $B^2 - 4AC$ the "discriminant". To get around this, we denote the determinant of the matrix by $\mathrm{disc}\, q$ and call it the **discriminant** of q, whereas we denote $B^2 - 4AC$ by $\Delta(q)$ and call it the **Discriminant** (i.e., with a capital D). Note that

$$\Delta(q) = -4 \,\mathrm{disc}(q).$$

Theorem 15.10. *(Hagedorn) Let $n \in \mathbb{Z}^+$ and let p be an odd prime. If $(\frac{-n}{p}) = 1$, then there exist $k, x, y \in \mathbb{Z}$ such that $x^2 + ny^2 = kp$ and*

$$1 \leq k \leq \lfloor \frac{4\sqrt{n}}{\pi} \rfloor.$$

Proof. The argument follows the $n = 1$ case rather closely. By assumption $-n$ is a square modulo p, so there exists $u \in \mathbb{Z}$ with $u^2 \equiv -n \pmod p$. Let

$$M := \begin{bmatrix} p & u \\ 0 & 1 \end{bmatrix}.$$

We have $\det(M) = p$, so $\Lambda := M\mathbb{Z}^2$ defines a lattice in $\mathbb{R}^2$ with

$$\mathrm{Covol}(\Lambda) = \det(M)\,\mathrm{Covol}(\mathbb{Z}^2) = p.$$

If $(t_1, t_2) \in \mathbb{Z}^2$ and $(x_1, x_2)^t = M(t_1, t_2)^t$, then

$$x_1^2 + nx_2^2 = (t_1 p + t_2 u)^2 + n t_2^2 \equiv (u^2+n) t_2^2 \equiv 0 \pmod p.$$

For $R > 0$, let

$$\Omega_R = \{(x,y) \in \mathbb{R}^2 \mid x^2 + ny^2 \leq R^2\}.$$

Then Ω_R is a compact, symmetric convex body with

$$\operatorname{Vol}\Omega_R = \frac{1}{\sqrt{n}}\operatorname{Vol}B_2 = \frac{\pi R^2}{\sqrt{n}},$$

so by Minkowski's Convex Body Theorem we have $\Lambda^\bullet \cap \Omega_R \neq \varnothing$ when

$$\operatorname{Vol}\Omega_R = \frac{\pi R^2}{\sqrt{n}} \geq 2^2 \operatorname{Covol}\Lambda = 4p,$$

so when

$$R^2 \geq \frac{4\sqrt{n}}{\pi}p.$$

If $v = (x, y) \in \Lambda^\bullet \cap \Omega_R$, then $(x, y) \in \mathbb{Z}^2$,

$$q(x, y) = x^2 + ny^2 \equiv 0 \pmod{p}$$

and

$$0 < q(x, y) \leq \frac{4\sqrt{n}}{\pi}p.$$

The result follows, since $q(x, y)$, being a positive integer multiple of p which is at most $\frac{4\sqrt{n}}{\pi}p$, must in fact be at most $\lfloor\frac{4\sqrt{n}}{\pi}\rfloor p$. □

We can slightly improve Hagedorn's result by using the sharp value of the lattice constant of B_2, namely $\Delta(B_2) = \frac{\sqrt{3}}{2}$ and the fact that the unique critical lattice for B_2 is the root lattice A_2. By definition of the lattice constant we may take

$$R^2 \geq \Delta(B_2)^{\frac{-2}{2}}(\operatorname{disc} q)^{\frac{1}{2}}(\operatorname{Covol}\Lambda)^{2/2} = \frac{2}{\sqrt{3}} \cdot \sqrt{n}p.$$

In other words, in the above result we may replace the constant $\frac{4}{\pi} \approx 1.273$ with $\frac{2}{\sqrt{3}} \approx 1.1547$, a small improvement! Moreover, let A be the linear transformation $(x, y) \mapsto (x, \frac{y}{\sqrt{n}})$, which maps the R-ball to the level set Ω_R. Then if the lattice $A\Lambda$ is *not* homothetic (i.e., equal up to dilation) to the root lattice A_2, then we may take the inequality to be strict.

Exercise: Show that for all n, p as above the lattice $A\Lambda$ is *not* homothetic to A_2.

Thus we get:

Theorem 15.11. *Under the hypotheses of Theorem 15.10, there are $x, y, k \in \mathbb{Z}^+$ with*

$$x^2 + ny^2 = kp$$

and

$$1 \leq k \leq \lfloor 2\sqrt{\frac{n}{3}}\rfloor,$$

$$k < 2\sqrt{\frac{n}{3}}.$$

The inequalities on k in Theorem 15.11 force $k = 1$ iff $1 \leq n \leq 3$, so in addition to Theorem 15.5 we get two more (very classical) representation theorems.

Corollary 15.12. *a) Suppose that p is a prime number such that $(\frac{-2}{p}) = 1$ – i.e., $p \equiv 1, 3 \pmod 8$. Then there are integers x, y such that $p = x^2 + 2y^2$.*
b) Suppose that p is an odd prime number such that $(\frac{-3}{p}) = 1$, i.e., $p \equiv 1 \pmod 3$. Then there are integers x, y such that $p = x^2 + 3y^2$.

Note that Hagedorn's Theorem is enough to deduce Corollary 15.12a) but to get Corollary 15.12b) in this way we really needed the full theory of lattice constants and critical lattices. This seems bad, because it suggests that we are stuck when $n \geq 4$. However, this is very far from the case!

Example: Consider x^2+3y^2 again. Using Hagedorn's bound we see that if $(\frac{-3}{p}) = 1$ then either $x^2 + 3y^2 = p$ or $x^2 + 3y^2 = 2p$. However it turns out that the second alternative can be easily **ruled out** using elementary congruence arguments. To wit: as above, by Quadratic Reciprocity, $(\frac{-3}{p}) = 1 \iff p \equiv 1 \pmod 3$. Reducing $x^2 + 3y^2 = 2p$ modulo 3 gives $x^2 \equiv 2 \pmod 3$, a blatant contradiction.

In fact one can take matters **much further**: the main result of [Ha11] is a description of which primes p (with $\gcd(p, 2n) = 1$) are represented by $x^2 + ny^2$ for 65 different positive integer values of n, the largest such being 1865. In most of these cases the form of the result is mildly different: one finds that there are **auxiliary congruence conditions** necessary (and, for these 65 forms, sufficient!) for p to be of the form $x^2 + ny^2$. Let us look at the simplest example of this.

Example: Consider $p = x^2 + 5y^2$. As ever, reducing modulo 5 shows that $(\frac{-5}{p}) = 1$. But now suppose reduce modulo 4: we get $p \equiv x^2 + 5y^2 \equiv x^2 + y^2 \pmod 4$, and as we have seen, 3 is not a sum of two squares modulo 4. Thus we get the *auxiliary congruence* $p \equiv 1 \pmod 4$. A little elementary work with congruences show that these two congruence conditions taken together amount to requiring $p \equiv 1, 9 \pmod{20}$. Miraculously, these easy necessary conditions are sufficient.

Theorem 15.13. *Let p be an odd prime number such that $(\frac{-5}{p}) = 1$.*
a) The following are equivalent:
(i) There are $x, y \in \mathbb{Z}$ such that $x^2 + 5y^2 = p$.
(ii) $p \equiv 1 \pmod 4$.
b) The following are equivalent:
(i) There are $x, y \in \mathbb{Z}$ such that $x^2 + 5y^2 = 2p$.
(ii) $p \equiv 3 \pmod 4$.

Proof. Step 1: Since $(\frac{-5}{p}) = 1$ and $\lfloor 2\sqrt{\frac{5}{3}} \rfloor = 2$, Theorem 15.11 tells us that there are integers x, y such that *either*

$$x^2 + 5y^2 = p$$

or

$$x^2 + 5y^2 = 2p.$$

Our task is now to show that the first alternative holds iff $p \equiv 1 \pmod 4$ and the second alternative holds if $p \equiv 3 \pmod 4$.
Step 2: Observe that $1 = (\frac{-5}{p}) = (\frac{-1}{p})(\frac{5}{p})$. Therefore, if $p \equiv 1 \pmod 4$, then $(\frac{-1}{p}) = (\frac{5}{p}) = 1$, whereas if $p \equiv 3 \pmod 4$ then $(\frac{-1}{p}) = (\frac{5}{p}) = -1$.

Case 1: $p \equiv 1 \pmod 4$. Then by Quadratic Reciprocity, $(\frac{5}{p}) = (\frac{p}{5}) = 1$, so p is a square modulo 5. In this case we can rule out $x^2 + 5y^2 = 2p$ by reducing modulo 5: we get $x^2 \equiv 2p \pmod 5$. But since $p \equiv y^2 \pmod 5$, this gives $(\frac{x}{y})^2 \equiv 2 \pmod 5$, i.e., 2 is a square mod 5: which it isn't! Therefore we must have $x^2 + y^2 = p$.
Case 2: $p \equiv 3 \pmod 4$. In this case Quadratic Reciprocity gives $-1 = (\frac{5}{p}) = (\frac{p}{5})$, so p is *not* a square modulo 5. Therefore if $x^2 + 5y^2 = p$ then reducing modulo 5 gives a contradiction, so we must have $x^2 + 5y^2 = 2p$. □

This was a rather innocuous case. To derive the congruence conditions under which a prime p is of the form, say, $x^2 + 1848y^2$ takes rather more work.

The quadratic forms treated by Hagedorn are precisely the known **principal** positive definite binary forms of Discriminant $-4n$ such that the Picard group of the imaginary quadratic order $\mathbb{Z}[\sqrt{-n}]$ has exponent dividing 2. Specifically, there are 65 discriminants, as follows:

$$\Delta = 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\ 12,\ 13,\ 15,\ 16,\ 18,\ 21,\ 22,\ 24,\ 25,\ 28,\ 30,\ 33,$$
$$37, 40,\ 42,\ 45,\ 48,\ 57,\ 58,\ 60,\ 70,\ 72,\ 78,\ 85,\ 88,\ 93,\ 102,\ 105,\ 112,\ 120,\ 130,\ 133,$$
$$165, 168,\ 177,\ 190,\ 210,\ 232,\ 240,\ 253,\ 273,\ 280,\ 312,\ 330,\ 345,\ 357,\ 385,\ 408, 462,$$
$$520,\ 760,\ 840,\ 1320,\ 1365,\ 1848.$$

A similar case is left open by Hagedorn's work: namely the principle positive definite binary quadratic forms of Discriminant $\Delta \equiv 1 \pmod 4$ such that the Picard group of $\mathbb{Z}[\frac{1+\sqrt{\Delta}}{2}]$ has exponent dividing 2. More precisely, for $\Delta \equiv 1 \pmod 4$ the principal form of Discriminant Δ is

$$q_1(x, y) = x^2 + xy + \frac{1 - \Delta}{4} y^2,$$

and the list of idoneal $\Delta \equiv 1 \pmod 4$ is as follows:

First the squarefree discriminants:

$$\Delta = -3, -7, -11, -15, -19, -35, -43, -51, -67, -91, -115, -123, -163, -187,$$
$$-195, -235, -267, -403, -427, -435, -483, -555, -595, -627, -715, -795,$$
$$-1155, -1435, -1995, -3003, -3315.$$

Finally, the non-squarefree discriminants:

$$\Delta = -27, -75, -99, -147, -315.$$

There should be $65 + 31 + 5 = 101$ such discriminants altogether: please check!

In order to make progress on these nondiagonal forms we need a version of Theorem 15.11 which applies to not necessarily diagonal positive definite forms. The following is a reasonable first guess as to what GoN methods should yield.

Conjecture 15.14. *Let $q(x, y) = Ax^2 + Bxy + Cy^2$ be a primitive, positive definite integral quadratic form. Put $\Delta(q) = B^2 - 4AC$. Let p be an odd prime with $(\frac{\Delta(q)}{p}) = 1$. Then there exist $x, y \in \mathbb{Z}$ and $k \in \mathbb{Z}^+$ with*

$$q(x, y) = kp$$

and

$$1 \leq k \leq \lfloor 2\sqrt{\frac{4AC - B^2}{12}} \rfloor = \lfloor \sqrt{\frac{4\,\mathrm{disc}(q)}{3}} \rfloor.$$

Conjecture 15.14 would immediately solve the representation problem for $\Delta = -3, -7, -11$ since in these cases we would have $k = 1$. (And with any luck the larger values of Δ listed above can be treated in roughly the same way we handled $x^2 + 5y^2$ above, i.e., as Hagedorn does in [Ha11].)

Update: Conjecture 15.14 has been proven by Hans Parshall and the author.

Also one can try moving beyond idoneal quadratic forms – to see that there is some chance at saying *something* here, see [Ha11, p. 13, proof of Prop. 3]. In fact, though I don't have the time to properly elaborate on this at present, this is closely connected to the concept of **bi-idoneal form** of Jagy and Kaplansky [JK].

For later use, we include the following additional representation theorem.

Theorem 15.15. *A prime p different from 2 and 5 is of the form $2x^2 + 5y^2$ iff $(\frac{-10}{p}) = 1$ and $p \equiv 2, 3 \pmod 5$ iff $p \equiv 7, 13, 23, 27 \pmod{40}$.*

Proof. Necessity: Suppose $p = 2x^2 + 5y^2$. Reducing modulo p gives $2x^2 + 5y^2 \equiv 0 \pmod p$. If $x \equiv 0 \pmod p$, then $p = 2x^2 + 5y^2$ shows $y \equiv 0 \pmod p$ and thus $p = 2x^2 + 5y^2$ is divisible by p^2, a contradiction. Therefore $2, 5, x, y$ are all invertible modulo p, and $2x^2 + 5y^2 \equiv 0 \implies \frac{-5}{2} \equiv X^2 \pmod p$, which holds iff $-10 \equiv X^2 \pmod p$, i.e., iff $(\frac{-10}{p}) = 1$. Similarly, reducing modulo 5 gives $p \equiv 2x^2 \pmod 5$, so p is not a square modulo 5 and hence $p \equiv 2, 3 \pmod 5$. Thus

$$1 = \left(\frac{-10}{p}\right) = \left(\frac{-2}{p}\right)\left(\frac{5}{p}\right) = -\left(\frac{-2}{p}\right),$$

so $\left(\frac{-2}{p}\right) = -1$, and thus $p \equiv 5, 7 \pmod 8$. A Chinese Remainder Theorem calculation gives $p \equiv 2, 3 \pmod 5$ and $p \equiv 5, 7 \pmod 8 \iff p \equiv 7, 13, 23, 27 \pmod{40}$. Sufficiency: Suppose $p > 5$ satisfies the necessary congruence conditions. In particular p is prime to 10 and such that $(\frac{-10}{p}) = 1$, so by Thue's Lemma there are $x, y, k \in \mathbb{Z}$ with $2x^2 + 5y^2 = kp$ and $1 \leq k < |2| + |0| + |5|$, i.e., $1 \leq k \leq 6$.[14] If $k = 1$, we're done.
• Suppose $2x^2 + 5y^2 = 2p$. Reducing modulo 5 gives $p \equiv x^2 \pmod 5$, contradiction.
• Suppose $2x^2 + 5y^2 = 3p$. Reducing modulo 5 gives $p \equiv x^2 \pmod 5$, contradiction.
• Suppose $2x^2 + 5y^2 = 4p$. Then y is even, and reducing modulo 4 shows x is even. So we may put $x = 2X$, $y = 2Y$ to get $2X^2 + 5Y^2 = p$.
• Suppose $2x^2 + 5y^2 = 5p$. Put $x = 5X$ to get $10X^2 + y^2 = p$. Reducing mod 5 shows $(\frac{p}{5}) = 1$, contradiction.
• Suppose $2x^2 + 5y^2 = 6p$. Put $y = 2Y$ to get $x^2 + 10Y^2 = 3p$. Reducing modulo 3 gives $x^2 + 10Y^2 \equiv x^2 + Y^2 \equiv 0 \pmod 3$, which forces x and Y both to be divisible by 3, contradiction. □

[14] The bound of the (now proven) Conjecture 15.14 would give $1 \leq k \leq 3$, but it is not much more trouble to consider values of k up to 6, so we do so.

15.5. **The Four Squares Theorem.**

Lemma 15.16. *(Euler's Identity) For any integers $a_1, \ldots, a_4, b_1, \ldots, b_4$, we have*

$$(a_1^2 + a_2^2 + a_3^2 + a_4^2)(b_1^2 + b_2^2 + b_3^2 + b_4^2) = (a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4)^2 +$$
$$(a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3)^2 + (a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2)^2 + (a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1)^2.$$

Proof. Again we apply Littlewood's Principle. □

Thus the set of sums of four integer squares is closed under multiplication. Since $1 = 1^2 + 0^2 + 0^2 + 0^2$ is a sum of four squares, it suffices to show that each prime p is a sum of four squares. Since $2 = 1^2 + 1^2 + 0^2 + 0^2$, we may assume $p > 2$.

Lemma 15.17. *For a prime $p > 2$ and $a \in \mathbb{Z}$, there exist $r, s \in \mathbb{Z}$ such that*

$$r^2 + s^2 \equiv a \pmod{p}.$$

Proof. There are $\frac{p-1}{2}$ nonzero squares mod p and hence $\frac{p-1}{2} + 1 = \frac{p+1}{2}$ squares mod p. Rewrite the congruence as $r^2 \equiv a - s^2 \pmod{p}$. Since the map $\mathbb{F}_p \to \mathbb{F}_p$ given by $t \mapsto a - t$ is an injection, as x ranges over all elements of $\mathbb{F}_p$ both the left and right hand sides take $\frac{p+1}{2}$ distinct values. Since $\frac{p+1}{2} + \frac{p+1}{2} > p$, these subsets cannot be disjoint, and any common value gives a solution to the congruence. □

Theorem 15.18. *(Lagrange) Every positive integer is a sum of four integral squares.*

Proof. By Lemma 15.17, there are $r, s \in \mathbb{Z}$ such that $r^2 + s^2 + 1 \equiv 0 \pmod{p}$. Define

$$M = \begin{bmatrix} p & 0 & r & s \\ 0 & p & s & -r \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

We have $\det(M) = p^2$, so $\Lambda := M\mathbb{Z}^4$ defines a lattice in $\mathbb{R}^4$ with

$$\mathrm{Vol}(\Lambda) = \det(M)\,\mathrm{Covol}(\mathbb{Z}^4) = p^2.$$

If $(t_1, t_2, t_3, t_4) \in \mathbb{Z}^4$ and $(x_1, x_2, x_3, x_4) := M(t_1, t_2, t_3, t_4)$ then

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = (pt_1 + rt_3 + st_4)^2 + (pt_2 + st_3 - rt_4)^2 + t_3^2 + t_4^2$$
$$\equiv\ t_3^2(r^2 + s^2 + 1) + t_4^2(r^2 + s^2 + 1)\ \equiv 0 \pmod{p}.$$

Now let

$$\Omega = B_0(\sqrt{2p}) = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1^2 + x_2^2 + x_3^2 + x_4^2 < 2p\}$$

be the open ball of radius $\sqrt{2p}$ about the origin in $\mathbb{R}^4$. Using Lemma **??** we have

$$\mathrm{Vol}(\Omega) = \frac{\pi^2}{2}(\sqrt{2p})^4 = 2\pi^2 p^2 > 16p^2 = 2^4\,\mathrm{Covol}\,\Lambda,$$

so by Minkowski's Theorem Mark II there exists $(x_1, \ldots, x_4) \in \Lambda$ with

$$0 < x_1^2 + x_2^2 + x_3^2 + x_4^2 < 2p.$$

Since $p \mid x_1^2 + x_2^2 + x_3^2 + x_4^2$, the only possible conclusion is

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = p.$$

□

15.5.1. *A Linear Forms Approach to the Four Squares Theorem.* In their paper [BR51], Brauer and Reynolds give a variant of the above proof which is interesting in that it replaces the appeal to Minkowski's Convex Body Theorem with the more elementary Brauer-Reynolds Theorem on linear forms. As above, we put $q(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 + x_3^2 + x_4^2$ and show that q $\mathbb{Z}$-represents every odd prime p. And again, by Lemma 15.17, there are $a, b \in \mathbb{Z}$ such that $a^2+b^2+1 \equiv 0 \pmod{p}$. Now consider the system of linear equations

$$x_1 \equiv ax_3 + bx_4 \pmod{p},$$

$$x_2 \equiv bx_3 - ax_4 \pmod{p}.$$

Here we have $r = 2$ equations in $s = 4$ unknowns. For $1 \leq i \leq 4$, take $\lambda_i = \sqrt{p} + \epsilon$. Then $\prod_{i=1}^s \lambda_i > p^2$, so we get $v = (x_1, x_2, x_3, x_4) \in \mathbb{Z}^4$ satisfying the above congruences and having $|x_i| \leq \sqrt{p} + \epsilon$. Thus of course we also have $|x_i| \leq \lfloor \sqrt{p} + \epsilon \rfloor$, and since $\sqrt{p}$ is not an integer, for sufficiently small ϵ we have $|x_i| \leq \lfloor \sqrt{p} + \epsilon \rfloor < \sqrt{p}$. Moreover,

$$x_1^2 + x_2^2 \equiv (a^2 + b^2)x_3^2 + (a^2 + b^2)x_4^2 \equiv -x_3^2 - x_4^2 \pmod{p},$$

so $x_1^2 + x_2^2 + x_3^2 + x_4^2 = kp$ for $k \in \mathbb{Z}^+$. Further, $x_1^2 + x_2^2 + x_3^2 + x_4^2 < 4\sqrt{p}^2 = 4p$, so $k \in \{1, 2, 3\}$.

If $k = 1$, great. Suppose $k = 2$, so

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 2p.$$

Without loss of generality $x_1 \equiv x_2 \pmod{2}$ and $x_3 \equiv x_4 \pmod{2}$, so

$$\left(\frac{x_1 + x_2}{2}\right)^2 + \left(\frac{x_1 - x_2}{2}\right)^2 + \left(\frac{x_3 + x_4}{2}\right)^2 + \left(\frac{x_3 - x_4}{2}\right)^2 = p.$$

Suppose $k = 3$, so

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 3p.$$

Without loss of generality $3 \mid x_1$, and by adjusting the signs on x_2, x_3, x_4 if necessary, we assume $x_2 \equiv x_3 \equiv x_4 \pmod{3}$. Then

$$\left(\frac{x_2 + x_3 + x_4}{3}\right)^2 + \left(\frac{x_1 + x_3 - x_4}{3}\right)^2 + \left(\frac{x_1 - x_2 + x_4}{3}\right)^2 + \left(\frac{x_1 + x_2 - x_3}{3}\right)^2 = p.$$

15.5.2. *A Variant.*

Theorem 15.19. *The integral quadratic form*

$$q(v) = x^2 + y^2 + z^2 + 4w^2$$

is positive universal.

Proof. Step 1: Suppose n is not divisible by 4. By the Four Squares Theorem there exist $x, y, z, w \in \mathbb{Z}$ such that $n = x^2 + y^2 + z^2 + w^2$. Since $4 \nmid n$, x, y, z, w cannot all be odd. Without loss of generality $w = 2W$ for $W \in \mathbb{Z}$ and thus

$$n = x^2 + y^2 + z^2 + (2W)^2 = x^2 + y^2 + z^2 + 4W^2.$$

Step 2: Any $n \in \mathbb{Z}^+$ may be written as $n = 4^a m$ with $m \in \mathbb{Z}$, m not divisible by 4. By Step 1, there are x, y, z, w such that $m = x^2 + y^2 + z^2 + 4w^2$, and thus

$$n = 4^a m = (2^a x)^2 + (2^a y)^2 + (2^a z)^2 + 4(2^a w)^2.$$

□

15.6. **The Quadratic Form $x_1^2 + ax_2^2 + bx_3^2 + abx_4^2$.**

The following theorem summarizes our application of the Convex Body Theorem to positive definite forms and the computation of the Hermite constant γ_4.

Theorem 15.20. *Let $q(x, y, z, w)$ be a positive definite real quaternary quadratic form. For positive R, let*

$$\Omega_R = \{x \in \mathbb{R}^4 \mid q(x) \leq R^2\}.$$

a) We have $\mathrm{Vol}(\Omega_R) = \frac{\pi^2 R^4}{2\sqrt{\mathrm{disc}\, q}}$.
b) By Minkowski's Convex Body Theorem, for any lattice $\Lambda \subset \mathbb{R}^4$, there exists $v \in \Lambda^\bullet$ with $q(v) \leq \frac{4\sqrt{2}}{\pi}(\mathrm{disc}\, q)^{\frac{1}{4}}\sqrt{\mathrm{Covol}\,\Lambda}$.
c) Since $\gamma_4 = \sqrt{2}$, for $\Lambda \subset \mathbb{R}^4$, there is $v \in \Lambda^\bullet$ with $q(v) \leq \sqrt{2}|\mathrm{disc}\, q|^{\frac{1}{4}}\sqrt{\mathrm{Covol}\,\Lambda}$.
d) Suppose the lattice Λ is given as $A\mathbb{Z}^4$ for $A \in M_4(\mathbb{R})$, and let q_A be the quadratic form $x \mapsto q(Ax)$. Then unless q_A is H-equivalent to

$$q_4 = x^2 + xz + y^2 - yz + z^2 - zw + w^2,$$

there exists $v \in \Lambda^\bullet$ with $q(v) < \sqrt{2}|\mathrm{disc}\, q|^{\frac{1}{4}}\sqrt{\mathrm{Covol}\,\Lambda}$.

Lemma 15.21. *(Twisted Euler Identity) Let $a, b, x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4 \in \mathbb{R}$. Then:*

$$(x_1^2 + ax_2^2 + bx_3^2 + abx_4^2)(y_1^2 + ay_2^2 + by_3^2 + aby_4^2) = (x_1y_1 - ax_2y_2 - bx_3y_3 - abx_4y_4)^2$$
$$+a(x_1y_2 + x_2y_1 + bx_3y_4 - bx_4y_3)^2 + b(x_1y_3 - ax_2y_4 + x_3y_1 + ax_4y_2)^2$$
$$+ab(x_1y_4 + x_2y_3 - x_3y_2 + x_4y_1)^2.$$

Proof. As usual, Littlewood's Principle suffices. □

The statement and proof given above are rather disingenuous: one does not simply pluck identities like this from thin air. I used the following MAGMA code (a trivial variant of code supplied to me by Jim Stankewicz) to find the above identity:

```
> K<a,b> := FunctionField(Rationals(),2);
L<x1,x2,x3,x4,y1,y2,y3,y4> := FunctionField(K,8);
> Q<i,j,k> := QuaternionAlgebra<L|-a,-b>;
> alpha := x1 + i*x2 + j*x3 + k*x4;
> beta := y1 + i*y2 + j*y3 + k*y4;
> alpha*beta;
(x1*y1 - a*x2*y2 - b*x3*y3 - a*b*x4*y4) + (x1*y2 + x2*y1 + b*x3*y4 - b*x4*y3)*i
+ (x1*y3 - a*x2*y4 + x3*y1 + a*x4*y2)*j + (x1*y4 + x2*y3 - x3*y2 + x4*y1)*k

> Norm(alpha);
x1^2 + a*x2^2 + b*x3^2 + a*b*x4^2
> Norm(beta);
y1^2 + a*y2^2 + b*y3^2 + a*b*y4^2
> Norm(alpha)*Norm(beta) - Norm(alpha*beta);
0
```

Question 15.22. *According to Ramanujan's theorem, pairs $(a, b) \in \mathbb{Z}^2$ for which $q_{a,b}(x) = x_1^2 + ax_2^2 + bx_3^2 + abx_4^2$ is positive universal are:*

$$(a, b) = (1, 1),\ (1, 2),\ (1, 3),\ (2, 2),\ (2, 3),\ (2, 4),\ (2, 5).$$

We have seen the very classical GoN argument that leads to a proof of the positive universality for $(1,1)$*. How many of the other* 6 *quaternary forms listed above can be shown to be positive universal by similar GoN methods?*

Theorem 15.23. *The integral form* $x^2+y^2+2z^2+2w^2$ *is positive universal.*

Proof. (Sketch) We apply the method for $q_{1,1}$ and the sharp bound for the lattice constant of B^4 – including showing that the forms in question are *not homothetic* to the critical form. Details still be to written by one of the students in the group! □

Although it is interesting and instructive to see that the sharp bound for $\Delta(B_4)$ gives the result to us with no additional work, it is also instructive – even more so, perhaps – to see that one can get away with the weaker Minkowski bound using some elementary descent arguments. So here is a second proof.

Proof. Step 1: Certainly $x^2+y^2+2z^2+2w^2$ $\mathbb{Z}$-represents 1 and 2, so by Lemma 15.21 it suffices to deal with the case of an odd prime p. Moreover, if $p \equiv 1 \pmod 4$, then by XX $p = x^2+y^2+2\cdot 0^2+2\cdot 0^2$, so we may assume $p \equiv 3 \pmod 4$. Applying the Minkowski bound as in the proof of Theorem XX, we get integers k, x, y, z, w with $1 \leq k \leq \lfloor \frac{8}{\pi} \rfloor = 2$ such that

$$x^2+y^2+2z^2+2w^2 = kp.$$

If $k=1$, we're done, so suppose $x^2+y^2+2z^2+2w^2 = 2p$. Then $x \equiv y \pmod 2$. Case 1: x and y are both even. So we may take $x = 2X$, $y = 2Y$ to get

$$2X^2+2Y^2+z^2+w^2 = p.$$

Case 2: x and y are both odd. Then

$$p = \frac{1}{2}(x^2+y^2)+z^2+w^2 = \left(\frac{x+y}{2}\right)^2+\left(\frac{x-y}{2}\right)^2+z^2+w^2 = X^2+Y^2+z^2+w^2.$$

Since $p \equiv 3 \pmod 4$, exactly 3 of X, Y, z, w are odd: without loss of generality suppose y and z are odd. Then

$$p = X^2+Y^2+2\left(\frac{z+w}{2}\right)^2+2\left(\frac{z-w}{2}\right)^2 = X^2+Y^2+2Z^2+2W^2.$$

□

Theorem 15.24. *The integral quadratic form*

$$q(v) = x^2+y^2+2z^2+8w^2$$

is positive universal.

Proof. Step 0: We may assume n is squarefree, so in particular $n \not\equiv 0 \pmod 4$.
Step 1: We claim that every $n \equiv 3 \pmod 4$ is $\mathbb{Z}$-represented by q. Indeed, by Theorem 15.23 there are $x, y, z, w \in \mathbb{Z}$ such that

$$n = x^2+y^2+2z^2+2w^2. \tag{28}$$

If w is even, we may substitute $w = 2W$ to get

$$n = x^2+y^2+2z^2+8W^2,$$

and similarly if z is even. Thus we may assume z, w are both odd. Reducing (28) modulo 4 gives $n \equiv x^2+y^2 \pmod 4$, so $n \not\equiv 3 \pmod 4$.
Step 2: Suppose n_1 and n_2 are odd positive integers both represented by q. We

claim that $n_1 n_2$ is also represented by q.
Indeed, if

$$n_1 = x_1^2 + x_2^2 + 2x_3^2 + 2(2x_4)^2,\ n_2 = y_1^2 + y_2^2 + 2y_3^2 + 2(2y_4)^2,$$

then by Lemma 15.21 we have

$$n_1 n_2 = z_1^2 + z_2^2 + 2z_3^2 + 2(2x_1y_4 + x_2y_3 - x_3y_2 + 2x_4y_1)^2. \tag{29}$$

with $z_1, z_2, z_3 \in \mathbb{Z}$. Equation (29) exhibits $n_1 n_2$ in the form $q(v)$ iff $x_2y_3 - x_3y_2$ is even. Now if n_1 is odd, then $x_1^2 + x_2^2$ is odd and thus exactly one of x_1, x_2 is even. By interchanging x_1 and x_2 if necessary, we may assume that x_2 is even. In exactly the same way we may assume that y_2 is even and thus that $x_2y_3 - x_3y_2$ is even.
Step 3: Every odd $n \in \mathbb{Z}^+$ is $\mathbb{Z}$-represented by q. Indeed, by Step 2 it is enough to show that every odd prime number p is $\mathbb{Z}$-represented by q. If $p \equiv 1 \pmod 4$, then already $p = x_1^2 + x_2^2$, whereas if $p \equiv 3 \pmod 4$ then q $\mathbb{Z}$-represents p by Step 1.
Step 4: Suppose $n = 2n' \equiv 2 \pmod 4$. Since n' is odd, by Step 3, there are integers y_1, y_2, y_3, y_4, with $y_2 = 2Y_2$, such that $n' = y_1^2 + y_2^2 + 2y_3^2 + 2(2y_4)^2$. Then

$$\begin{aligned} n = 2 \cdot n' &= (0^2 + 0^2 + 2 \cdot 1^2 + 2(2 \cdot 0)^2)(y_1^2 + y_2^2 + 2y_3^2 + 2(2y_4)^2) \\ &= z_1^2 + z_2^2 + z_3^2 + 2(-y_2)^2 = z_1^2 + z_2^2 + z_3^2 + 8Y_2^2. \end{aligned}$$

□

Theorem 15.25. *The integral form $q = x^2 + y^2 + 3z^2 + 3w^2$ is positive universal.*

Proof. Clearly q $\mathbb{Z}$-represents 1 and 2, so by Lemma 15.21 it is enough to show that q represents every odd prime p. As above, there is an index p^2 sublattice Λ_p of $\mathbb{Z}^4$ such that for all $v \in \Lambda_p$, $q(v) \equiv 0 \pmod p$. Since $\operatorname{disc}(q) = 9$, by Theorem 15.20 there exists $(x, y, z, w) \in \mathbb{Z}^4$ and $k \in \mathbb{Z}$, $0 < k \leq \lfloor \sqrt{2 \cdot 3} \rfloor = 2$ such that

$$x^2 + y^2 + 3z^2 + 3w^2 = kp.$$

In other words, either $x^2 + y^2 = 3z^2 + 3w^2 = p$ – and we're done – or $x^2 + y^2 + 3z^2 + 3w^2 = 2p$. If so,

$$0 \equiv 2p \equiv x^2 + y^2 + 3z^2 + 3w^2 \equiv x + y + z + w \pmod 2.$$

Case 1: $x + y$, $z + w$ are both even. Then $\frac{x \pm y}{2}, \frac{z \pm w}{2} \in \mathbb{Z}$, so

$$\left(\frac{x+y}{2}\right)^2 + \left(\frac{x-y}{2}\right)^2 + 3\left(\frac{z+w}{2}\right)^2 + 3\left(\frac{z-w}{2}\right)^2 = \frac{2p}{2} = p,$$

and we have found a $\mathbb{Z}$-representation of p.
Case 2: $x + y$ and $z + w$ are both odd. Without loss of generality x and z are odd and z and w are even, so

$$2p \equiv x^2 + y^2 + 3z^2 + 3w^2 \equiv 1 + 3 \equiv 0 \pmod 4,$$

and thus p is even: contradiction! □

Remark: Here we used the sharp bound for $\Delta(B^4)$, a nontrivial theorem of Korkine-Zolotareff (the proof of which does not appear in these notes). It would be nice if we could get away with the bound afforded by the convex body theorem. In this case, this involves entertaining also $k = 3$. Someone should try this!

Theorem 15.26. *The integral form $q = x^2 + 2y^2 + 2z^2 + 4w^2$ is positive universal.*

Proof. The same opening strategy as in the proof of Theorem 15.26 reduces us to showing that every odd prime p is $\mathbb{Z}$-represented by q and shows that there are integers x, y, z, w, k with $x^2 + 2y^2 + 2z^2 + 4w^2 = kp$ and $1 \leq k \leq \lfloor \sqrt{2 \cdot 4} \rfloor = 2$. If $k = 1$ we're done, so suppose

$$x^2 + 2y^2 + 2z^2 + 4w^2 = 2p.$$

Then x is even, so taking $x = 2X$ and simplifying gives

$$2X^2 + y^2 + z^2 + 2w^2 = p.$$

Since p is odd, so is $y^2 + z^2$, so exactly one of y and z is even – without loss of generality, suppose y is even. Thus we may write $y = 2Y$ to get

$$z^2 + 2X^2 + 2w^2 + 4Y^2 = p.$$

□

Remark: As above, to use MCBT one needs also to look at $k = 3$. Try it!

Theorem 15.27. *The integral form $q = x^2 + 2y^2 + 3z^2 + 6w^2$ is positive universal.*

Proof. (Mordell-Hicks-Thompson-Walters)
Step 1: First consider the identity

$$x^2 + (y + z + w)^2 + (y - z - w)^2 + (z - 2w)^2 = x^2 + 2y^2 + 3z^2 + 6w^2,$$

and the inverse identity

$$x^2 + y^2 + z^2 + w^2 = x^2 + \left(\frac{y+z}{2}\right)^2 + \left(\frac{y-z+w}{3}\right)^2 + \left(\frac{y-z-2w}{6}\right)^2.$$

Let $n \in \mathbb{Z}^+$. By Theorem 15.18, there are $x, y, z, w \in \mathbb{Z}$ with $n = x^2 + y^2 + z^2 + w^2$, so

$$n = x^2 + \left(\frac{y+z}{2}\right)^2 + \left(\frac{y-z+w}{3}\right)^2 + \left(\frac{y-z-2w}{6}\right)^2.$$

This gives an integral representation of n by q provided all the following congruence conditions are satisfied:

$$y + z \equiv 0 \pmod 2,$$

$$y - z + w \equiv 0 \pmod 3,$$

$$y - z - 2w \equiv 0 \pmod 6.$$

Step 2: By Lemma 15.21, it suffices to show that q integrally represents 1 and all prime numbers. Certainly q $\mathbb{Z}$-represents 1, 2 and 3, so it suffices to show that q $\mathbb{Z}$-represents all primes $p > 3$. Thus $p \equiv 1, 5 \pmod 6$. Since $x^2 + 3y^2$ already represents all primes $p \equiv 1 \pmod 3$, we may assume $p \equiv 5 \pmod 6$.
Step 3: Let $p \equiv 5 \pmod 6$ be a prime. As above there are $x, y, z, w \in \mathbb{Z}$ such that

$$p = x^2 + y^2 + z^2 + w^2. \tag{30}$$

Now we consider (30) as a congruence modulo 6. The squares modulo 6 are $0, 1, 3, 4$. Without loss of generality we may take the congruence classes of x, y, z, w in non-decreasing order, and then there are four ways for $x^2 + y^2 + z^2 + w^2 \equiv 5 \pmod 6$:

$$0 + 0 + 1 + 4 \equiv 0 + 1 + 1 + 3 \equiv 0 + 3 + 4 + 4 \equiv 1 + 3 + 3 + 4 \equiv 5 \pmod 6.$$

Case 1: By adjusting the signs on x, y, z, w we may assume $x \equiv y \equiv 0 \pmod 6$, $z \equiv 1 \pmod 6$, $w \equiv 2 \pmod 6$ and write

$$p = y^2 + 2\left(\frac{w+x}{2}\right)^2 + 3\left(\frac{w-x+z}{3}\right)^2 + 6\left(\frac{w-x-2z}{6}\right)^2.$$

Case 2: We may assume $x \equiv 0 \pmod 6$, $y \equiv z \equiv 1 \pmod 6$, $w \equiv 3 \pmod 6$. Then

$$p = x^2 + 2\left(\frac{y+z}{2}\right)^2 + 3\left(\frac{y-z+w}{3}\right)^2 + 6\left(\frac{y-z-2w}{6}\right)^2.$$

Case 3: We may assume $x \equiv 0 \pmod 6$, $y \equiv 3 \pmod 6$, $z \equiv w \equiv 4 \pmod 6$. Then

$$p = x^2 + 2\left(\frac{z+w}{2}\right)^2 + 3\left(\frac{z-w+y}{3}\right)^2 + 6\left(\frac{z-w-2y}{6}\right)^2.$$

Case 4: We may assume $x \equiv 1 \pmod 6$, $y \equiv z \equiv 3 \pmod 6$, $w \equiv 4 \pmod 6$. Then

$$p = z^2 + 2\left(\frac{y+x}{2}\right)^2 + 3\left(\frac{y-x+w}{3}\right)^2 + 6\left(\frac{y-x-2w}{6}\right)^2.$$

□

Theorem 15.28. *The integral form $q = x^2 + 2y^2 + 4z^2 + 8w^2$ is positive universal.*

Proof. Certainly q $\mathbb{Z}$-represents 1 and 2. By Lemma 15.21 it is enough to show that q $\mathbb{Z}$-represents every odd prime. Moreover, by our work on binary forms, we know that every $p \equiv 1 \pmod 4$ is $\mathbb{Z}$-represented by $x^2 + 4z^2$, so we may assume $p \equiv 3 \pmod 4$. By Theorem 15.26 there are $x, y, z, w \in \mathbb{Z}$ such that

$$p = x^2 + 2y^2 + 2z^2 + 4w^2. \tag{31}$$

If y is even, we may put $y = 2Y$ to get $p = x^2 + 2z^2 + 4w^2 + 8Y^2$, and similarly if z is even. Finally, suppose that y and z are both odd. Certainly x is odd, so reducing (31) modulo 4 gives

$$p \equiv x^2 + 2y^2 + 2z^2 + 4w^2 \equiv 1 + 2 + 2 \equiv 1 \pmod 4.$$

□

15.7. **Beyond Universal Forms.**

A positive definite integral quadratic form $q(x)$ is called **almost universal** if it $\mathbb{Z}$-represents all but finitely many positive integers. The recent paper [BO09] gives definitive results on almost universal forms. Here of course we are interested in results which can be proved by our elementary GoN methods.

For instance Halmos showed [Ha38] that there are precisely 88 diagonal positive definite quadratic forms $(a, b, c, d) = ax^2 + by^2 + cz^2 + dw^2$ which represent all positive integers with exactly one exception. By the Ramanujan-Dickson Theorem this exceptional integer must be at most 15. Here is Halmos's list:

Forms representing all positive integers except 1:

$$(2, 2, 3, 4),\ (2, 3, 4, 5),\ (2, 3, 4, 8).$$

Forms representing all positive integers except 2:

$$(1, 3, 3, 5),\ (1, 3, 5, 6).$$

Forms representing all positive integers except 3:

$$(1,1,4,5),\ (1,1,4,6),\ (1,1,5,5),\ (1,1,5,6),\ (1,1,5,10),\ (1,1,5,11),$$

$$(1,1,6,7),\ (1,1,6,8),\ (1,1,6,10),\ (1,1,6,11).$$

Exercise: Why are there no forms representing all positive integers except 4? Except 8? Except 9? Except 12?
Forms representing all positive integers except 5:

$$(1,2,6,6),\ (1,2,6,10),\ (1,2,6,11),\ (1,2,6,12),\ (1,2,6,13),$$

$$(1,2,7,8),\ (1,2,7,10),\ (1,2,7,11),\ (1,2,7,12),\ (1,2,7,13).$$

Forms representing all positive integers except 6:

$$(1,1,3,7),\ (1,1,3,8),\ (1,1,3,10),\ (1,1,3,11),\ (1,1,3,13),\ (1,1,3,14),\ (1,1,3,15).$$

Forms representing all positive integers except 7:

$$(1,1,1,9),\ (1,1,1,10),\ (1,1,1,12),\ (1,1,1,14),\ (1,1,1,15),$$

$$(1,2,2,9),\ (1,2,2,10),\ (1,2,2,12),\ (1,2,2,14),\ (1,2,2,15).$$

Forms representing all positive integers except 10:

$$(1,2,3,11),\ (1,2,3,12),\ (1,2,3,13),\ (1,2,3,15),\ (1,2,3,17),(1,2,3,19),$$

$$(1,2,3,20),\ (1,2,3,21),\ (1,2,3,22),\ (1,2,3,23),\ (1,2,3,24),\ (1,2,3,5),\ (1,2,3,26),$$

$$(1,2,5,11),\ (1,2,5,12),\ (1,2,5,13),\ (1,2,5,14).$$

Exercise: Why are there no forms representing all positive integers except 11? Except 13? (The previous exercise has an easy answer. At the moment it is not clear to me that this one does.)

Forms representing all positive integers except 14:

$$(1,1,2,15),\ (1,1,2,17),\ (1,1,2,18),\ (1,1,2,19),\ (1,1,2,20),\ (1,1,2,21),$$

$$(1,1,2,22),\ (1,1,2,23),\ (1,1,2,24),\ (1,1,2,25),\ (1,1,2,27),\ (1,1,2,28),$$

$$(1,1,2,29),\ (1,1,2,30),\ (1,2,4,15),\ (1,2,4,17),\ (1,2,4,18),\ (1,2,4,19),$$

$$(1,2,4,20),\ (1,2,4,21),\ (1,2,4,22),\ (1,2,4,23),\ (1,2,4,24),\ (1,2,4,25),$$

$$(1,2,4,27),\ (1,2,4,28),\ (1,2,4,29),\ (1,2,4,30).$$

Forms representing all integers except 15:

$$(1,2,5,5).$$

Of the above, the easiest to deal with are those of square discriminant:

$$(1,1,1,9),\ (1,1,5,5),\ (1,2,6,12),\ (1,2,2,9),\ (1,2,3,24).$$

15.8. **Wójcik's Proof of the Three Squares Theorem.**

Theorem 15.29. *(Legendre-Gauss) A positive integer is a sum of three squares of integers iff it is* not *of the form* $4^a(8k+7)$.

Lemma 15.30. *Let* n *be an integer of the form* $4^a(8k+7)$ *for some* $a \in \mathbb{N}$, $k \in \mathbb{Z}$. *Then* n *is* not *the sum of three rational squares.*

Proof. Step 0: Suppose $4^a(8k+7)$ is a sum of three rational squares. We may take our rational numbers to have a common deminator $d > 0$ and thus

$$\left(\frac{x}{d}\right)^2 + \left(\frac{y}{d}\right)^2 + \left(\frac{z}{d}\right)^2 = 4^a(8k+7).$$

Clearing denominators, we get

$$x^2 + y^2 + z^2 = d^2 4^a(8k+7).$$

Write $d = 2^b d'$ with d' odd. Since $1^2, 3^2, 5^2, 7^2 \equiv 1 \pmod 8$, $d'^2 \equiv 1 \pmod 8$, so

$$d^2 4^a(8k+7) = (2^b)^2(d'^2)4^a(8k+7) = 4^{a+b}(8k'+7).$$

In other words, to show that no integer of the form $4^a(8k+7)$ is a sum of 3 rational squares, it suffices to show that no integer of the form $4^a(8k+7)$ is a sum of three integral squares. So let us now show this.
Step 1: We observe that $x^2 + y^2 + z^2 \equiv 7 \pmod 8$ has no solutions. Indeed, since the squares mod 8 are $0, 1, 4$, this is a quick mental calculation. (In particular this disposes of the $a = 0$ case.)
Step 2: we observe that if $n \equiv 0, 4 \pmod 8$ then the congruence

$$x^2 + y^2 + z^2 \equiv n \pmod 8$$

has no *primitive solutions*, i.e., no solutions in which at least one of x, y, z is odd. Indeed, since the squares mod 8 are $0, 1, 4$, so in particular the only odd square is 1. Since 4 and 0 are both even, if x, y, z are not all even, then exactly one two of them must be odd, say x and y, so $x^2 \equiv y^2 \equiv 1 \pmod 8$ and thus $z^2 \equiv 4 - 2 \pmod 8$ or $z^2 \equiv 8 - 2 \pmod 8$, and neither 2 nor 6 is a square modulo 8.
Step 3: Now suppose that there are integers x, y, z such that $x^2+y^2+z^2 = 4^a(8k+7)$. If $a = 0$ then by Step 1 reducing modulo 8 gives a contradiction. If $a = 1$, then $4^a(8k+7) \equiv 4 \pmod 8$, so by Step 2 any representation $x^2 + y^2 + z^2 = 4(8k+7)$ must have x, y, z all even, and then dividing by 4 gives $(\frac{x}{2})^2+(\frac{y}{2})^2+(\frac{z}{2})^2 = (8k+7)$, a contradiction. If $a \geq 2$, then $4^a(8k+7) \equiv 0 \pmod 8$, and again by Step 2 in any representation $x^2 + y^2 + z^2 = 4^a(8k+7)$ we must have x, y, z all even. Thus writing $x = 2X$, $y = 2Y$, $z = 2Z$ we get an integer representation $X^2 + Y^2 + Z^2 = 4^{a-1}(8k+7)$. We may continue in this way until we get a representation of $4(8k+7)$ as a sum of three integral squares, which we have just seen is impossible. □

Lemma 15.31. *Suppose that every squarefree positive integer* $n \not\equiv 7 \pmod 8$ *is a sum of three integral squares. Then every positive integer* $n \neq 4^a(8k+7)$ *is a sum of three integral squares.*

Proof. Let n be a positive integer which is *not* of the form $4^a(8k+7)$. Write n as $n = 2^a n_1^2 n_2$, where $a \geq 0$, n_1 is odd and n_2 is odd and squarefree.
Case 1: $0 \leq a \leq 1$, $n_2 \not\equiv 7 \pmod 8$. Then $2^a n_2$ is squarefree and not 7 (mod 8), so by assumption there exist $x, y, z \in \mathbb{Z}$ such that $x^2 + y^2 + z^2 = 2^a n_2$, and thus $(n_1 x)^2 + (n_1 y)^2 + (n_1 z)^2 = 2^a n_1^2 n_2 = n$.
Case 2: $n_2 \not\equiv 7 \pmod 8$. In such a case n is of the form $(2^b)^2$ times an integer n

of the type considered in Case 1. Since such an integer n is a sum of three integral squares, so is any square times n.
Case 3: $n_2 \equiv 7 \pmod 8$. For n not to be of the form $4^a(8k+7)$, the power of a must be odd; in other words, we may write n as a square times $2n_2$ where n_2 is squarefree and of the form $8k+7$. Thus $2n_2$ is squarefree and not of the form $8k+7$, so by assumption $2n_2$ is a sum of three squares, hence so is n. □

Lemma 15.32. *Let $x, y, z \in \mathbb{Q}$ be such that $x^2+y^2+z^2 \in \mathbb{Z}$. Then there exist $a, b, c \in \mathbb{Q}$ such that $a^2+b^2+c^2 = 1$ and $ax+by+cz \in \mathbb{Z}$.*

Proof. Let $x = \frac{x_1}{d},\ y = \frac{y_1}{d},\ z = \frac{z_1}{d}$, with $\gcd(x_1y_1z_1, d) = 1$. Let

$$\tilde{\Lambda} = \{(u+tx, v+ty, w+tz) \mid x, y, t \in \mathbb{Z}, t \in [0, d-1]\}.$$

and

$$\Lambda = \{(u+tx, v+ty, w+tz) \in \tilde{\Lambda} \mid ux+vy+wz \in \mathbb{Z}\}.$$

Then $\tilde{\Lambda} = \mathbb{Z}^3 + \langle (x,y,z) \rangle$, so $[\tilde{\Lambda} : \mathbb{Z}^3] = d$. Further, $ux+vy+wz \in \mathbb{Z} \iff ux_1+vy_1+wz_1 \equiv 0 \pmod d$, so $[\tilde{\Lambda} : \Lambda] \leq d$. It follows that $\operatorname{Covol} \Lambda \leq 1$.
Let

$$\Omega = \{(a,b,c) \in \mathbb{R}^3 \mid a^2+b^2+c^2 < 2\}.$$

Then

$$\operatorname{Vol}(\Omega) = \frac{4\pi}{3} \cdot (\sqrt{2})^3 > 8 > 8 \operatorname{Covol} M,$$

so by MCBT, there exists $(a,b,c) \in M$ such that

$$(a,b,c) = (u+tx, v+ty, w+tz),\ 0 < a^2+b^2+c^2 < 2.$$

Since

$$a^2+b^2+c^2 = u^2+v^2+w^2+2t(ux+vy+wz)+t^2(x^2+y^2+z^2) \in \mathbb{Z},$$

we must have

$$a^2+b^2+c^2 = 1$$

and

$$ax+by+cz = ux+vy+wx+t(x^2+y^2+z^2) \in \mathbb{Z}.$$

□

Lemma 15.33. *The integral quadratic form $q(x,y,z) = x^2+y^2+z^2$ is an ADC-form: every integer which is $\mathbb{Q}$-represented by q is $\mathbb{Z}$-represented by q.*

Proof. Let $x, y, y \in \mathbb{Q}$ be such that $x^2+y^2+z^2 \in \mathbb{Z}$. Let $a, b, c \in \mathbb{Q}$ be as in Lemma 15.32. We may assume $b^2+c^2 \neq 0$. Then

$$x^2+y^2+z^2 = (ax+by+cz)^2 + U^2 + V^2,$$

where

$$U = bx - \frac{ab+c^2}{b^2+c^2}y + \frac{-abc+bc}{b^2+c^2}z,$$

$$V = cx + \frac{-abc+bc}{b^2+c^2}y - \frac{ac^2+b^2}{b^2+c^2}z.$$

The integer U^2+V^2 is a sum of two squares of rational numbers hence a sum of two squares of integers. □

Lemma 15.34. *Let $m \in \mathbb{Z}^+$, $n \equiv 3 \pmod 8$, and write $m = p_1 \cdots p_r$. Then the number of i such that $p_i \equiv 3, 5 \pmod 8$ is even.*

Exercise: Prove Lemma 15.34. (Suggestion: use the Jacobi symbol $\left(\frac{-2}{m}\right)$.)

Proposition 15.35. *Let n be a squarefree integer, $n \not\equiv 7 \pmod 8$. Then n is a sum of three rational squares.*

Proof. To fix ideas we will first give the argument under certain additional congruence conditions and then explain how to modify it to deal with the other cases. Filling in the details for these latter cases is a good exercise for the interested reader.
Case 1: Let us suppose that $m = p_1 \cdots p_r$ is squarefree and $m \equiv 1 \pmod 4$. Thus each p_i is odd and the number of $p_i \equiv 3 \pmod 4$ is even. By Dirichlet's Theorem on Primes in Arithmetic Progressions, there is a prime number q such that
• $\left(\frac{q}{p_i}\right) = \left(\frac{-1}{p_i}\right)$ for all $1 \leq i \leq p_i$ and
• $q \equiv 1 \pmod 4$.
(Indeed, each of the first conditions restricts q to a nonempty set of congruence classes modulo the distinct odd primes p_i, whereas the last condition is a condition modulo a power of 2. By the Chinese Remainder Theorem this amounts to a set of congruence conditions modulo $4p_1 \cdots p_r$ and all of the resulting congruence classes are relatively prime to $4p_1 \cdots p_r$, so Dirichlet's Theorem applies.)
It follows that for all $1 \leq i \leq r$,

$$\left(\frac{-q}{p_i}\right) = \left(\frac{-1}{p_i}\right)\left(\frac{q}{p_i}\right) = 1,$$

and

$$\left(\frac{m}{q}\right) = \left(\frac{p_1}{q}\right)\cdots\left(\frac{p_r}{q}\right) = \left(\frac{q}{p_1}\right)\cdots\left(\frac{q}{p_r}\right) = \left(\frac{-1}{p_1}\right)\cdots\left(\frac{-1}{p_r}\right) = 1.$$

The last equality holds because the number of factors of -1 is the number of primes $p_i \equiv 3 \pmod 4$, which as observed above is an even number.
since $-q$ is a square modulo each of the distinct primes p_i, by the Chinese Remainder Theorem it is also a square modulo $m = p_1 \cdots p_r$. Therefore by the Chinese Remainder Theorem there is an integer x such that

$$x^2 \equiv -q \pmod m$$
$$x^2 \equiv m \pmod q.$$

But according to Legendre's Theorem, these are precisely the congruence conditions necessary and sufficient for the homogeneous equation

$$qu^2 + z^2 - mt^2 = 0$$

to have a solution in integers (u, z, t), not all zero. Indeed, we must have $t \neq 0$, for otherwise $qu^2 + z^2 = 0 \implies u = z = 0$. Moreover, since $q \equiv 1 \pmod 4$, by Fermat's Two Squares Theorem there are $x, y \in \mathbb{Z}$ such that $qu^2 = x^2 + y^2$. Therefore

$$mt^2 - z^2 = qu^2 = x^2 + y^2,$$

so

$$m = \left(\frac{x}{t}\right)^2 + \left(\frac{y}{t}\right)^2 + \left(\frac{z}{t}\right)^2$$

and m is a sum of three rational squares, completing the proof in this case.
Case 2: Suppose $m = 2m_1 = 2p_1 \cdots p_r$ with $m_1 = p_1 \cdots p_r$ squarefree and odd. In this case we may proceed exactly as above, except that we require $q \equiv 1 \pmod 8$.
Case 3: Suppose $m = p_1 \cdots p_r$ is squarefree and $m \equiv 3 \pmod 8$. By Lemma 15.34, the number of prime divisors p_i of m which are either 5 or 7 modulo 8 is even. By

Dirichlet's Theorem there exists a prime q such that
• $\left(\frac{q}{p_i}\right) = \left(\frac{-2}{p_i}\right)$ for all $1 \leq i \leq p_i$ and
• $q \equiv 5 \pmod 8$.
It follows that for all $1 \leq i \leq r$,

$$\left(\frac{-2q}{p_i}\right) = \left(\frac{-2}{p_i}\right)\left(\frac{q}{p_i}\right) = 1,$$

and

$$\left(\frac{m}{q}\right) = \left(\frac{p_1}{q}\right)\cdots\left(\frac{p_r}{q}\right) = \left(\frac{q}{p_1}\right)\cdots\left(\frac{q}{p_r}\right) = \left(\frac{-2}{p_1}\right)\cdots\left(\frac{-2}{p_r}\right) = 1.$$

The last equality holds because the number of factors of -1 is the number of primes $p_i \equiv 5, 7 \pmod 8$, which as observed above is an even number.
Therefore there is an integer x such that

$$x^2 \equiv -2q \pmod m$$

$$x^2 \equiv m \pmod q,$$

so by Legendre's Theorem the equation

$$2qu^2 + z^2 - mt^2 = 0$$

has a solution in integers (u, z, t) with $t \neq 0$. Since $q \equiv 1 \pmod 4$, there are $x, y \in \mathbb{Z}$ such that $2qu^2 = x^2 + y^2$, so

$$mt^2 - z^2 = 2qu^2 = x^2 + y^2,$$

and thus once again

$$m = \left(\frac{x}{t}\right)^2 + \left(\frac{y}{t}\right)^2 + \left(\frac{z}{t}\right)^2.$$

□

Remark 9. *Let us emphasize that the GoN contribution of Wójcik's argument is precisely to establish that $q = x^2 + y^2 + z^2$ is an **ADC form**: that is, for every $n \in \mathbb{Z}$, if there exist rational numbers (x, y, z) such that $q(x, y, z) = n$, then there exist integers (x, y, z) such that $q(x, y, z) = n$. My own recent work on quadratic forms has centered around the study of ADC-forms: see* [Cl11].

*The name "ADC-form" comes from the following obervation of Aubry and Davenport-Cassels: suppose an anisotropic integral quadratic form $q(x) = q(x_1, \ldots, x_n)$ has the following **Euclidean** property: for every $x = (x_1, \ldots, x_n) \in \mathbb{Q}^n$, there exists $y = (y_1, \ldots, y_n) \in \mathbb{Z}$ such that $|q(x - y)| < 1$. Then it is necessarily an ADC-form. In particular – and indeed this was exactly the case treated by Aubry – the form $q = x_1^2 + x_2^2 + x_3^2$ is a Euclidean form, as is easily seen: approximating a rational triple $x = (x_1, x_2, x_3)$ by a nearest integer triple $y = (y_1, y_2, y_3)$ we clearly have*

$$|q(x - y)| \leq \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{3}{4} < 1.$$

This provides a "geometric argument", of some sort, that the sum of three squares form is an ADC form, which to me seems simpler than Wójcik's. W.C. Jagy and I have classified all positive definite ternary Euclidean forms over $\mathbb{Z}$ and also all positive definite ADC forms [ADCII]*, and there are about ten times as many ADC forms as Euclidean forms. I would be very interested to know whether GoN methods*

such as Wójcik's above can be used to show the ADC property for any ternary forms which are not *Euclidean.*

15.9. **Ankeny's Proof of the Three Squares Theorem.**

In view of what has already been done – especially Lemma 15.31 – we may safely construe the "three squares theorem" to be the assertion that every positive squarefree integer $m \not\equiv 7 \pmod 8$ is a sum of three integral squares. The argument will require different computations depending on the congruence class of m modulo 8. We will treat first the case $m \equiv 3 \pmod 8$ and then afterwards discuss modifications necessary to treat $m \equiv 1, 2, 5, 6 \pmod 8$.

Case 1: Let $m \equiv 3 \pmod 8$ be positive and squarefree: put

$$m = p_1 \cdots p_r.$$

By Dirichlet's Theorem there is a prime q such that

$$\left(\frac{-2q}{p_j}\right) = 1, \ 1 \leq j \leq r, \tag{32}$$

$$q \equiv 1 \pmod 4. \tag{33}$$

By (32) and (33) we have

$$1 = \prod_{j=1}^{r} \left(\frac{-2q}{p_j}\right) = \prod_{j=1}^{r} \left(\frac{-2}{p_h}\right)\left(\frac{q}{p_j}\right)$$

$$= \left(\frac{-2}{m}\right) \prod_{j=1}^{r} \left(\frac{p_j}{q}\right) = \left(\frac{-2}{m}\right)\left(\frac{m}{q}\right) = \left(\frac{-2}{m}\right)\left(\frac{-m}{q}\right) = \left(\frac{-m}{q}\right),$$

since $m \equiv 3 \pmod 8$. Thus there is an integer b such that $b^2 \equiv -m \pmod q$. Replacing b by $b + m$ if needed, we may assume b is odd. There is $h_1 \in \mathbb{Z}$ such that

$$b^2 - qh_1 = -m. \tag{34}$$

Considering (34) modulo 4 shows $h_1 = 4h$ for $h \in \mathbb{Z}$, so

$$b^2 - 4qh = -m. \tag{35}$$

Since m is squarefree, (32) implies there is $t \in \mathbb{Z}$ such that

$$t^2 \equiv \frac{-1}{2q} \pmod m. \tag{36}$$

Now we consider the ellipsoid

$$\Omega = \{(R, S, T) \in R^3 \mid R^2 + S^2 + T^2 < 2m\} \tag{37}$$

where

$$R = 2tqx + tby + mz, \ S = \sqrt{2q}x + \frac{b}{\sqrt{2q}}y, \ T = \sqrt{\frac{m}{2q}}y.$$

Thus $(R, S, T)^t = M_A(x, y, z)^t$ where

$$M_A = \begin{bmatrix} 2tq & tb & m \\ \sqrt{2q} & \frac{b}{\sqrt{2q}} & 0 \\ 0 & \sqrt{\frac{m}{2q}} & 0 \end{bmatrix}.$$

By Lemma 9.5 we have $\operatorname{Vol}\Omega = \frac{4\pi}{3}(2m)^{3/2}$. Since $\det M_A = m^{3/2}$, the ellipsoid $M_A^{-1}\Phi$ (as Ankeny says, the body "in (x,y,z)-space") has volume

$$\frac{2^{7/2}\pi}{3} \approx 11.847687835 > 8.$$

So by Minkowski's Convex Body Theorem, there is $(x_1, y_1, z_1) \in (\mathbb{Z}^3)^\bullet$ satisfying (37). Let $(R_1, S_1, T_1)^t = M_A(x_1, y_1, z_1)^t$. Now for some unpleasant algebra:

$$R_1^2 + S_1^2 + T_1^2 = (2tqx_1 + tby_1 + mz_1)^2 + \left((\sqrt{2q}x_1 + \frac{b}{\sqrt{2q}}y_1\right)^2 + \left(\sqrt{\frac{m}{2q}}y_1\right)^2$$

$$\equiv t^2(2qx_1 + by_1)^2 + \frac{1}{2q}(2qx_1 + by_1)^2 \equiv 0 \pmod{m}.$$

Moreover

$$R_1^2 + S_1^2 + T_1^2 = R_1^2 + \left((\sqrt{2q}x_1 + \frac{b}{\sqrt{2q}}y_1\right)^2 + \left(\sqrt{\frac{m}{2q}}y_1\right)^2$$

$$= R_1^2 + \frac{1}{2q}(2qx_1 + by_1)^2 + \frac{m}{2q}y_1^2 = R_1^2 + 2(qx_1^2 + bx_1y_1 + hy_1^2).$$

Put

$$v = qx_1^2 + bx_1y_1 + hy_1^2. \tag{38}$$

Certainly $v \in \mathbb{Z}$. Moreover, the binary quadratic form $qx^2 + bxy + hy^2$ has determinant $b^2 - 4qh = -m < 0$ so is positive definite. Thus $v \in \mathbb{N}$, and $v = 0$ only if $x_1 = y_1 = 0$, so $z_1 \neq 0$ and upon plugging back in we would get

$$R^2 + S^2 + T^2 = m^2z_1^2 < 2m,$$

and thus $mz_1^2 < 2$. Since $m \geq 3$, this is a contradiction and thus $v \in \mathbb{Z}^+$. Also $R_1 \in \mathbb{Z}$ and the unpleasant algebra gives $m \mid R_1^2 + 2v$, whereas by (37) we have $0 < R_1^2 + 2v < m$. It follows that

$$R_1^2 + 2v = m. \tag{39}$$

It suffices to show that $2v = X^2 + Y^2$ is a sum of two integer squares, for then $m = R_1^2 + X^2 + Y^2$. By Fermat's Theorem it is sufficient to show that for all $p > 2$ with $\operatorname{ord}_p(v)$ odd, we have $p \equiv 1 \pmod 4$. So let $p > 2$ be such that $\operatorname{ord}_p(v)$ is odd.

∘ Suppose $p \nmid m$. Then by (39) we have

$$\left(\frac{m}{p}\right) = 1. \tag{40}$$

By (38) we get (recall that $b^2 + m = 4qh$):

$$4qv = (2qx_1 + by_1)^2 + my_1^2.$$

If $p \mid q$ then $\left(\frac{-m}{p}\right) = 1$.

If $p \nmid q$ then $\operatorname{ord}_p((2qx_1 + by_1^2 + my_1^2))$ is odd, which implies $\left(\frac{-m}{p}\right) = 1$. Either way

$$\left(\frac{-m}{p}\right) = 1,$$

which combined with (40) yields $\left(\frac{-1}{p}\right) = 1$, i.e., $p \equiv 1 \pmod 4$.

◦ Suppose $p \mid v$, $p \mid m$. Then

$$m = R_1^2 + 2v = R_1^2 + \frac{1}{2q}((2qx_1 + by_1)^2 + my_1^2), \tag{41}$$

so $p \mid R_1$ and $p \mid (2qx_1 + by_1)$. Since m is squarefree we have $\gcd(\frac{m}{p}, p) = 1$ and then dividing both sides of (41) by p we get

$$\frac{1}{2q}\frac{m}{p}y_1^2 \equiv \frac{m}{p} \pmod p$$

and thus

$$y_1^2 \equiv 2q \pmod p,$$

so $\left(\frac{2q}{p}\right) = 1$. Combining this with (32) we get $\left(\frac{-1}{p}\right) = 1$, i.e., $p \equiv 1 \pmod 4$.

This shows that for every odd prime p with $\mathrm{ord}_p(v)$ is odd we have $p \equiv 1 \pmod 4$. The same holds for $2v$, so we get that there are $X, Y \in \mathbb{Z}$ such that $2v = X^2 + Y^2$. Substituting back in (39) we get

$$m = R_1^2 + 2v = R_1^2 + X^2 + Y^2,$$

establishing the result in this case.

Case 2: Let $m \equiv 1, 2, 5, 6 \pmod 8$ be positive and squarefree. We alter the above argument as follows: choose a prime $q \equiv 1 \pmod 4$ such that $\left(\frac{-q}{p_j}\right) = 1$ for all odd prime divisors p_j of m and such that, if m is even we have

$$m = 2m_1, \left(\frac{-2}{q}\right) = (-1)^{\frac{m_1-1}{2}}, \ t^2 \equiv \frac{-1}{q} \pmod{p_j},$$

$$t \text{ odd}, b^2 - qh = -m$$

and

$$R = tqx + tby + mz, \ S = \sqrt{q}x + \frac{b}{\sqrt{q}}y, \ T = \sqrt{\frac{m}{q}}y.$$

The proof then proceeds identically to that of Case 1.

15.10. **Mordell's Proof of the Three Squares Theorem.**

As above, it will suffice to show that every positive squarefree integer $n \not\equiv 7 \pmod 8$ is a sum of three integral squares. It is this result for which Mordell gives a new proof, which is in some ways a streamlining of Ankeny's proof.

Step 1: Let $a, b, h \in \mathbb{Z}$. Put

$$m = ab - h^2 \tag{42}$$

and

$$\varphi(x, y) = ax^2 + 2hxy + by^2,$$

so

$$\mathrm{disc}\, \varphi = ab - h^2 = m.$$

Further, let $A, B \in \mathbb{Z}$ and define a ternary quadratic form $f(x, y, z) \in \mathbb{Q}[x, y, z]$ by

$$mf(x, y, z) = (Ax + By + mz)^2 + \varphi(x, y). \tag{43}$$

Notice that when $a, b, m > 0$, φ is positive definite and thus so is f. The symmetric matrix corresponding to the right hand side of (43) is

$$\begin{bmatrix} A^2+a & AB+h & Am \\ AB+h & B^2+b & Bm \\ Am & Bm & m^2 \end{bmatrix}, \tag{44}$$

so

$$m^3 \det f = \det(mf) = m^2(ab-h^2) = m^3$$

and thus

$$\det f = 1.$$

We claim that we can choose these parameter values so as to make f classically integral, i.e., such that the matrix of (44) has integral entries. This holds if $A, B, a, b, h, m \in \mathbb{Z}$ and

$$A^2 + a \equiv B^2 + b \equiv AB + h \equiv 0 \pmod{m}. \tag{45}$$

We will establish this in Step 2. Assuming it for now: since $\gamma_3 = 2^{1/3} < 2$, we have that f integrally represents 1. By Hermite's Lemma, there is a $\mathbb{Z}$-basis e_1, e_2, e_3 such that $f(e_1) = 1$. Moreover, being classically integral is coordinate-invariant – it means that the corresponding symmetric bilinear form $B_f = q(x+y) - q(x) - q(y)$ is defined over $\mathbb{Z}$ – so up to $\mathrm{GL}_3(\mathbb{Z})$-equivalence we have

$$f \sim x^2 + 2fxy + ny^2 + 2gxz + 2kyz + \ell z^2.$$

Because the coefficients of xy and xz are even, we can complete the square to eliminate them, getting

$$f \sim x^2 + n'y^2 + 2k'yz + \ell' z^2.$$

(Compare with Theorem 15.2!) The binary form $n'y^2 + 2k'yz + \ell z^2$ is positive definite and has discriminant 1 hence Discriminant $\Delta = -4$, and by Proposition 9.16 there is only one such form up to $\mathrm{SL}_2(\mathbb{Z})$-equivalence: x^2+y^2. (Hence *a fortiori* there is only one such form up to integral equivalence.) Thus altogether we find

$$f \sim x^2 + y^2 + z^2 =: g.$$

Since $m = f(0,0,1)$, $m \in f(\mathbb{Z}^3)$, hence also $m \in g(\mathbb{Z}^3)$.

Step 2: Let $m \in \mathbb{Z}^+$ be squarefree and $m \not\equiv 7 \pmod 8$. For $a \in \mathbb{Z}$, there are $b, h \in \mathbb{Z}$ satisfying (42) iff $-m$ is a square modulo a. We will take $a = \delta a_1$ with $\delta \in \{1, 2\}$ and a_1 an odd prime. In this case, (42) is solvable for $b, h \in \mathbb{Z}$ iff

$$\left(\frac{-m}{a_1}\right) = 1. \tag{46}$$

Suppose (46) holds. Since $b = \frac{h^2+m}{a}$, we can still find $b \in \mathbb{Z}$ which satisfies $ab - h^2 = m$ after adding any multiple of a to h, and thus – since $\gcd(a, m) = 1$ – we may assume $h \equiv 0 \pmod m$ hence also $b \equiv 0 \pmod m$. If we take $B \equiv 0 \pmod m$ then $B^2 + b \equiv AB + h \equiv 0 \pmod m$. This gives two of the conditions of (45); the last is $A^2 + a \equiv 0 \pmod m$. Such an $A \in \mathbb{Z}$ exists iff $-a$ is a square modulo m.

To sum up: fix $m \in \mathbb{Z}^+$. If we can find an odd prime number a_1 and $\delta \in \{1, 2\}$ such that for $a = \delta a_1$ we have that $-a$ is a square modulo m and that $-m$ is a square modulo a_1, then a is a sum of three integral squares. We have thus reduced to an exercise in quadratic reciprocity, which we solve by considering various cases.

Case 1: Suppose $m \equiv 1 \pmod 4$. We take $\delta = 1$ so $a_1 = a$. By Dirichlet's Theorem

on primes in arithmetic progressions there is a prime a with $a \equiv 1 \pmod 4$ and $a \equiv -1 \pmod m$. Then $-a$ is a square modulo m. Moreover

$$\left(\frac{-m}{a_1}\right) = \left(\frac{-m}{a}\right) = \left(\frac{m}{a}\right) = \left(\frac{a}{m}\right) = \left(\frac{-1}{m}\right) = 1.$$

Case 2: Suppose $m \equiv 2 \pmod 8$, so $m \equiv 2m_1$ with $m_1 \equiv 1 \pmod 4$. We take $\delta = 1$, so $a_1 = a$. By Dirichlet's Theorem there is a prime a with $a \equiv 1 \pmod 8$ and $a \equiv -1 \pmod m$. Then $-a$ is a square modulo m. Moreover

$$\left(\frac{-m}{a_1}\right) = \left(\frac{-m}{a}\right) = \left(\frac{m_1}{a}\right) = \left(\frac{a}{m_1}\right) = \left(\frac{-1}{m_1}\right) = 1.$$

Case 3: Suppose $m \equiv 6 \pmod 8$, so $m \equiv 2m_1$ with $m_1 \equiv 3 \pmod 4$. We take $\delta = 1$ so $a_1 = a$. By Dirichlet's Theorem there is a prime a with $a \equiv 3 \pmod 8$ and $a \equiv -1 \pmod{m_1}$. Then $-a$ is a square modulo m. Moreover

$$\left(\frac{-m}{a_1}\right) = \left(\frac{-m}{a}\right) = \left(\frac{m_1}{a}\right) = -\left(\frac{a}{m_1}\right) = \left(\frac{-a}{m_1}\right) = \left(\frac{1}{m_1}\right) = 1.$$

Case 4: Suppose $m \equiv 3 \pmod 8$. We take $\delta = 2$, so $a = 2a_1$. By Dirichlet's Theorem, there is a prime a_1 with $a_1 \equiv 1 \pmod 4$ and $a = 2a_1 \equiv -1 \pmod m$. Then $-a$ is a square modulo m. Moreover

$$\left(\frac{-m}{a_1}\right) = \left(\frac{m}{a_1}\right) = \left(\frac{a_1}{m}\right) = \left(\frac{-2a_1}{m}\right) = \left(\frac{1}{m}\right) = 1.$$

15.11. **Some applications of the Three Squares Theorem.**

Knowing which integers are represented by $x^2 + y^2 + z^2$ is a powerful weapon for analyzing representation of integers by certain quaternary quadratic forms.

Proposition 15.36. *The three squares theorem implies the four squares theorem.*

Proof. In order to show the Four Squares Theorem it suffices to show that every squarefree positive integer m is a sum of four integer squares. By the Three Squares Theorem, m is even a sum of three integer squares unless $m = 8k + 7$. But if $m = 8k+7$, then $m-1 = 8k+6$. Now $\mathrm{ord}_2(8k+6) = 1$, so $8k+6$ is not of the form $4^a(8k + 7)$, hence $8k + 6 = x^2 + y^2 + z^2$ and $m = 8k + 7 = x^2 + y^2 + z^2 + 1^2$. □

More generally:

Theorem 15.37. *For any $1 \leq d \leq 7$, the quadratic form $q = x^2 + y^2 + z^2 + dw^2$ integrally represents all positive integers.*

Proof. As above it is enough to show that q represents all squarefree positive integers. Moreover, if $m \neq 8k+7$ is a squarefree positive integer then m is represented already by $x^2 + y^2 + z^2$ so certainly by q. It remains to dispose of $m = 8k + 7$.
Case 1: Suppose $d = 1, 2, 4, 6$. Then $m - d \cdot 1^2 = m - d$ is:
• $m - 1 = 8k + 6$, if $d = 1$. This is a sum of 3 squares.
• $m - 2 = 8k + 5$, if $d = 2$. This is a sum of 3 squares.
• $m - 4 = 8k + 3$, if $d = 3$. This is a sum of 3 squares.
• $m - 5 = 8k + 2$, if $d = 5$. This is a sum of 3 squares.
• $m - 6 = 8k + 1$, if $d = 6$. This is a sum of 3 squares.
Case 2: If $d = 3$, then

$$m - d \cdot 2^2 = m - 12 = 8k - 5 = 8(k - 1) + 3.$$

Thus, so long as $m - 12$ is positive, it is a sum of three squares. We need to check separately that positive integers less than 12 are still represented by q, but this is easy: the only one which is not already a sum of 3 squares is $7 = 2^2+0^2+0^2+3\cdot 1^2$.
Case 3: If $d = 7$, then

$$m - d \cdot 2^2 = m - 28 = 8(k-3) + 5.$$

Thus, so long as $m - 28$ is positive, it is a sum of three squares. Again we must separately check that positive integers less than 28 are represented by q, and again this comes down to checking 7: $7 = 0^2 + 0^2 + 0^2 + 7 \cdot 1^2$. □

If we are looking for quaternary quadratic forms $q = x^2 + y^2 + z^2 + dw^2$ which represent *all* positive integers, then we have just found all of them: if $d > 7$, then such a q cannot integrally represent 7. Nevertheless we can still use the Gauss-Legendre Theorem to analyze these forms. For instance.

Proposition 15.38. *For a positive integer n, TFAE:*
(i) There are integers x, y, z, w such that $n = x^2 + y^2 + z^2 + 8w^2$.
(ii) $n \not\equiv 7 \pmod 8$.

Proof. (i) $\implies$ (ii): For any integers x, y, z, w, reducing $n = x^2 + y^2 + z^2 + 8w^2$ modulo 8 gives $n \equiv x^2 + y^2 + z^2 \pmod 8$, and we already know that this has no solutions when $n \equiv 7 \pmod 8$.
(ii) $\implies$ (i): Write $n = 2^a m$ with m odd. If m is not of the form $8k + 7$ then both m and $2m$ are sums of three integer squares, and since n is an even power of 2 times either m or $2m$, n must be a sum of three integer squares. So we are reduced to the case $n = 2^a(8k + 7)$ with $a \geq 1$. If $a = 1$ then $\mathrm{ord}_2(n) = 1$ and again n is a sum of three integer squares. Suppose $a = 2$, so $n = 32k + 28$ and thus $n - 8 \cdot 1^2 = 32k + 20 = 4(8k + 5)$ is of the form $x^2 + y^2 + z^2$ and thus $n = x^2 + y^2 + z^2 + 8w^2$. If $a \geq 3$ is odd, then n is a sum of three squares. If $a \geq 4$ is even, then $n = (2^{\frac{a-2}{2}})^2(4 \cdot (8k + 7))$ is a square times an integer represented by q, so n is also represented by q. □

Exercise: Prove or disprove the following claims:
a) If d is a positive integer which is not divisible by 8, then the quadratic form $x^2 + y^2 + z^2 + dw^2$ integrally represents all sufficiently large positive integers.
b) If $d = 8d'$ is a positive integer, then the quadratic form $x^2 + y^2 + z^2 + dw^2$ integrally represents all sufficiently large positive integers which are *not* $7 \pmod 8$.

15.12. **The Ramanujan-Dickson Ternary Forms.**

Consider the following seven positive definite ternary quadratic forms:

$$q_A(x, y, z) = x^2 + y^2 + z^2.$$
$$q_B(x, y, z) = x^2 + y^2 + 2z^2.$$
$$q_C(x, y, z) = x^2 + y^2 + 3z^2.$$
$$q_D(x, y, z) = x^2 + 2y^2 + 2z^2.$$
$$q_E(x, y, z) = x^2 + 2y^2 + 3z^2.$$
$$q_F(x, y, z) = x^2 + 2y^2 + 4z^2.$$
$$q_G(x, y, z) = x^2 + 2y^2 + 5z^2.$$

These forms are the key to the proof of the **Diagonal Fifteen Theorem** mentioned above. Indeed, above we used Theorem 15.29 to prove the universality of the forms $(1,1,1,d)$ for $1 \leq d \leq 7$. Ramanujan's proof of the universality of the other $54-7$ diagonal forms is similar, but also uses the forms q_B through q_G.

The Three Squares Theorem was proven by Legendre and Gauss at the beginning of the 19th century. The representation theorem for q_C was proved by Dirichlet in 1850. I don't know about the history of the representation theorems for the other five forms: Ramanujan states these results but does not prove or reference them in his paper. Concerning this, in the introduction to [Dic27], Dickson makes the following rather uncharitable comment "He gave no proofs for these forms and doubtless obtained his results empirically." Dickson takes it upon himself to give proofs of the other six representation theorems,[15] and in so doing completes the proof of the Diagonal Fifteen Theorem.

The evident challenge here is to give proofs of the other six representation theorems by GoN methods. In fact, for three of these forms this is not necessary, since the proofs that Dickson gives are by reducing them to q_A.

Theorem 15.39. *For $n \in \mathbb{Z}^+$, the following are equivalent:*
(i) n is not *of the form $2^{2a+1}(8k+7)$.*
(ii) n is integrally represented by $q_B = x^2+y^2+2z^2$.

Proof. [Dic27] (i) $\implies$ (ii): Suppose first that $n = 2k+1$ is odd. By Theorem 15.29, there exist $x,y,z \in \mathbb{Z}$ such that $x^2+y^2+z^2 = 4k+2 = 2n$. Reduction modulo 4 shows that exactly two of x,y,z are odd: say x,y are odd and $z = 2Z$ is even. Then $X = \frac{x+y}{2}$, $Y = \frac{x-y}{2} \in \mathbb{Z}$, so

$$2n = (X+Y)^2 + (X-Y)^2 + (2z)^2 = 2X^2 + 2Y^2 + 4Z^2,$$

and thus $n = X^2+Y^2+2Z^2$. Next suppose $m \neq 2^{2a}(8k+7)$. By Theorem 15.29, there exist $x,y,z \in \mathbb{Z}$ such that $X^2+Y^2+z^2 = m$, so

$$2m = 2X^2+2Y^2+2z^2 = (X+Y)^2+(X-Y)^2+2z^2,$$

and every even positive integer not of the form $2^{2a+1}(8k+7)$ is represented by q_B.
(ii) $\implies$ (i): Suppose $x^2+y^2+2z^2 = n$. We need to show $n \neq 2^{2a+1}(8k+7)$: this is clear if n is odd. Otherwise, if $n = 2m = x^2+y^2+2z^2$, then x and y have the same parity so we may put $X = \frac{x+y}{2}$, $Y = \frac{x-y}{2}$ to get $2m = (X+Y)^2+(X-Y)^2+2z^2 = 2X^2+2Y^2+2z^2$, so $m = X^2+Y^2+Z^2$. By Theorem 15.29 m is not of the form $2^{2a}(8k+7)$, so $n = 2m$ is not of the form $2^{2a+1}(8k+7)$. □

Theorem 15.40. *For $n \in \mathbb{Z}^+$, the following are equivalent:*
(i) n is not *of the form $4^a(8k+7)$.*
(ii) n is integrally represented by $q_D = x^2+2y^2+2z^2$.

Proof. [Dic27] (i) $\implies$ (ii): Suppose first that n is odd and not of the form $8k+7$. By Theorem 15.29 there are $x,y,z \in \mathbb{Z}$ such that $n = x^2+y^2+z^2$. Since n is odd, at least one of x,y,z is odd: without loss of generality suppose x is odd; then y

[15] Our labelling of the seven forms is the same as Dickson's, except we have interchanged q_B and q_C from [Dic27]. Our (lexicographic) ordering seems more straightforward and easier to remember.

and z have the same parity. Thus we may define $Y = \frac{y+z}{2}$, $Z = \frac{y-z}{2}$, so $y = Y + Z$, $z = Y - Z$, and then

$$n = x^2 + y^2 + z^2 = x^2 + (Y+Z)^2 + (Y-Z)^2 = x^2 + 2Y^2 + 2Z^2 = q_D(x, Y, Z).$$

Next suppose $n = 2r$ is not of the form $4^a(8k+7)$. Then r is not of the form $2^{2a+1}(8k+7)$, so by Theorem 15.39 there are $x, y, z \in \mathbb{Z}$ such that $r = x^2+y^2+2z^2$, and thus

$$n = 2r = 2x^2 + 2y^2 + 4z^2 = (2z)^2 + 2x^2 + 2y^2 = q_D(2z, x, y).$$

(ii) $\implies$ (i): $x^2 + 2y^2 + 2z^2 = x^2 + (y+z)^2 + (y-z)^2$. Apply Theorem 15.29. □

Theorem 15.41. *For $n \in \mathbb{Z}^+$, the following are equivalent:*
(i) n is not *of the form $2^{2a+1}(8k+7)$.*
(ii) n is integrally represented by $q_F = x^2 + 2y^2 + 4z^2$.

Proof. [Dic27] (i) $\implies$ (ii): Let n be a positive integer not of the form $2^{2a+1}(8k+7)$. Suppose first that n is odd, so by Theorem 15.39 there are $x, y, z \in \mathbb{Z}$ such that $n = x^2 + y^2 + 2z^2$. Then x and y have opposite parity, so without loss of generality $x = 2X$ is even, and thus $n = y^2+2z^2+4X^2 = q_F(y, z, X)$. Next suppose n is even, so $n = 2m$ with m not of the form $4^a(8k+7)$. By Theorem 15.40 there are $x, y, z \in \mathbb{Z}$ such that $m = x^2 + 2y^2 + 2z^2$, so $n = 2m = (2x)^2 + 2y^2 + 4z^2 = q_F(2x, y, z)$.
(ii) $\implies$ (i): Suppose $n = x^2 + 2y^2 + 4z^2$. If n is odd it is obviously not of the form $2^{2a+1}(8k+7)$, so suppose $n = 2m$ is even. Then $n = 2m = x^2 + 2y^2 + 4z^2$, so we may write $x = 2X$ to get

$$m = y^2 + 2z^2 + 2X^2.$$

By Theorem 15.40 $m \neq 4^a(8k+7)$, so $n = 2m \neq 2^{2a+1}(8k+7)$. □

The significance of the next result lies in the "global" method of proof. It uses a result – the Aubry-Davenport-Cassels Lemma – that we have not described here (but see e.g. [Cl11]), but I want to record it here and now because it was communicated to me orally by Allan Lacy, and I don't want to forget it.

Theorem 15.42. *The form $q_B = x^2 + y^2 + 2z^2$ is an ADC-form.*

Proof. (Lacy) The form q is **boundary-Euclidean**: for all $v \in \mathbb{Q}^3$, there exists $w \in \mathbb{Z}^3$ such that $q(v-w) < 1$ *unless* $v = (x, y, z)$ with $x, y, z \in \frac{1}{2} + \mathbb{Z}$, in which case the nearest integer approximation gives $q(v-w) = 1$. However, if $x = x_0 + \frac{1}{2}$, $y = y_0 + \frac{1}{2}$, $z = z_0 + \frac{1}{2}$ are half-integers such that

$$d = q(x, y, z) = (x_0 + \frac{1}{2})^2 + (y_0 + \frac{1}{2})^2 + 2(z_0 + \frac{1}{2})^2,$$

so

$$4d = (2x_0+1)^2 + (2y_0+1)^2 + 2(2z_0+1)^2 \equiv 4 \pmod 8,$$

so d is odd. But by Theorem 15.39, q $\mathbb{Z}$-represents d anyway, so the implication "q $\mathbb{Q}$-represents d $\implies$ q $\mathbb{Z}$-represents d" certainly holds for all odd d! □

Anyway, this leaves us with the task of proving representation theorems for q_C, q_E and q_G using GoN methods. Dickson proves them using **reduction theory**, which in truth is regarded as a main branch of GoN, but it is a tool that we have not yet used in our study of quadratic forms. Given that such proofs already exist, the open problem seems here is the existence of reduction-less proofs. For instance, someone should at least try to adapt Wójcik's proof of Theorem 15.29.

16. Applications of GoN: Isotropic Vectors for Quadratic Forms

16.1. **Cassels's Isotropy Theorem.**

For $v = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we put

$$||v|| = \max_i |x_i|.$$

Similarly, for a matrix $A = (a_{ij}) \in M_n(\mathbb{R})$, we put

$$||A|| = \max_{i,j} |a_{ij}|.$$

Finally, for an n-ary quadratic form q with coefficients in $\mathbb{R}$, we put

$$||q|| = ||A||,$$

where $A \in M_n(K)$ is the corresponding symmetric matrix, i.e., $q(t) = t^T At$.

Now let q be an integral n-ary quadratic form. An **isotropic vector** v for q is a nonzero vector $v \in \mathbb{Z}^n$ such that $q(v) = 0$. We say that q is **isotropic** if it has an isotropic vector; otherwise we say q is **anisotropic**.

Theorem 16.1. *(Cassels* [Ca55]*) Let q be an n-ary quadratic form with coefficients in $\mathbb{Z}$, $q \not\equiv 0$. If q is isotropic, there is an isotropic vector v for q with*

$$||v|| \leq (3||q||)^{\frac{n-1}{2}}.$$

Proof. Let $a = (a_1, \ldots, a_n) \in \mathbb{Z}^n$ be a nonzero anisotropic vector for q with $||a||$ minimal. By relabelling the variables if necessary, we may assume $||a|| = |a_1|$. Further, we may assume $|a_1| \geq 2$: if $||a|| = |a_1| = 1$, then $||a|| = 1 \leq \sqrt{3}^{\frac{n-1}{2}} \leq (3||q||)^{\frac{n-1}{2}}$. For $v, w \in \mathbb{Q}^n$, we define the associated bilinear form

$$\langle v, w \rangle = \frac{q(v+w) - q(v) - q(w)}{2}.$$

For all $v \in \mathbb{Q}^n$, $\langle v, v \rangle = q(v)$.
Step 1: We claim that for all $b \in \mathbb{Z}^n$ with $q(b) \neq 0$ and all $c \in K$,

$$||q||^{\frac{-1}{2}} \leq \sqrt{3}||b - ca||.$$

Let

$$a^* = q(b)a - 2\langle a, b \rangle b.$$

Then $a^* = q(b)\tau_b(a)$, where $\tau_b \in O(q)$ is reflection through the anisotropic vector b. It follows that since $q(a) = 0$, $q(a^*) = 0$. Or, if you like, just calculate directly:

$$q(a^*) = \langle a^*, a^* \rangle = \langle q(b)a - 2\langle a, b \rangle b, q(b)a - 2\langle a, b \rangle b \rangle$$
$$= q(b)^2\langle a, a \rangle - 4\langle a, b \rangle^2 q(b) + 4\langle a, b \rangle^2 q(b) = 0,$$

since $q(a) = 0$. By the minimality of a, we have

$$||a|| \leq ||a^*||. \tag{47}$$

Now put $d = b - ca$, so $b = d + ca$. Then

$$a^* = q(d + ca)a - 2\langle a, d + ca \rangle (d + ca)$$
$$= (q(d) + 2c\langle a, d \rangle + c^2 q(a))a - 2\langle a, d(d + ca)$$
$$= q(d)a - 2\langle a, d \rangle d.$$

Using the Archimedean property of the standard absolute value on $\mathbb{Z}$, we get

$$||a^*|| = ||q(d)a - 2\langle a, d\rangle d|| \leq |q(d)|||a|| + 2|\langle a, d\rangle|||d|| \leq 3||q||||a||||d||^2. \tag{48}$$

Combining (47) and (48) and dividing through by $||a||$, we get

$$1 \leq 3||q|| \cdot ||b - ca||^2,$$

or equivalently

$$||q||^{\frac{-1}{2}} \leq \sqrt{3}||b - ca||. \tag{49}$$

Step 2: We claim there is $b \in \mathbb{Z}^n$ and $c \in \mathbb{Q}$ with $q(b) \neq 0$ and

$$||b - ca|| \leq ||a||^{\frac{-1}{n-1}}. \tag{50}$$

Apply Corollary 13.4 with $n-1$ in place of n, $M = |a_1|$, $\theta_i = \frac{a_i}{a_1}$. Then there is $0 < b_1 < |a_1|$ and $b_2, \ldots, b_n \in \mathbb{Z}$ such that $|b_i - \frac{b_1}{a_1}a_i| \leq ||a||^{\frac{-1}{n-1}}$. If we take $c = \frac{b_1}{a_1}$ this defines $b \in \mathbb{Z}^n$ with $||b - ca|| \leq ||a||^{\frac{-1}{n-1}}$. Further, for all $2 \leq i \leq n$, we have

$$|b_i| \leq |\frac{b_1}{a_1}a_i| + |a_1|^{\frac{-1}{n-1}} < |b_1| + 1,$$

so $||b|| = |b_1| < |a_1| = ||a||$. By minimiality of $||a||$, this forces $q(b) \neq 0$.
Step 3: Combining (49) and (50) we get

$$||q||^{\frac{-1}{2}} \leq \sqrt{3}||a||^{\frac{-1}{n-1}},$$

or

$$||a|| \leq (3||q||)^{\frac{n-1}{2}}.$$

□

16.2. **Legendre's Theorem.**

Let $q(x, y, z)$ be a nondegenerate quadratic form over $\mathbb{Z}$. Is q isotropic? This depends only on the $\mathbb{Q}$-isomorphism class of q, so we may diagonalize and take

$$q(x, y, z) = ax_1^2 + bx_2^2 + cx_3^2 = 0,$$

with a, b, c nonzero *squarefree* integers. Now if a, b, c are all positive, then q is positive definite, hence certainly anisotropic over $\mathbb{Z}$; similarly q is anisotropic if a, b, c are all negative. Up to relabeling and multiplying through by -1, we may – and shall – assume that a is positive and b and c are negative.

Finally, we may reduce to the case in which a, b, c are coprime in pairs, or equivalently that abc is squarefree. We leave this as a simple but enlightening exercise for the reader. Thus we are led to consider the **Legendre equation**

$$ax^2 + by^2 + cz^2 = 0, \tag{51}$$

with $a > 0$, $b, c < 0$ and abc squarefree.

Thus we are led to consider the **Legendre equation**

$$ax^2 + by^2 + cz^2 = 0, \tag{52}$$

with $a > 0$, $b, c < 0$ and abc squarefree.

We claim that if $q(x, y, z)$ is isotropic, then $-bc$ is a square modulo a, $-ac$ is a square modulo b, and $-ab$ is a square modulo c. Indeed, suppose there are

$x, y, z \in \mathbb{Q}$, not all zero, such that $ax^2 + by^2 + cz^2 = 0$. By rescaling, we may assume that $(x, y, z) \in \mathbb{Z}^3$ and $\gcd(x, y, z) = 1$.

Let p be a prime dividing a. Reducing 52 modulo a gives

$$by^2 + cz^2 \equiv 0 \pmod{p}.$$

If y and z were both divisible by p, then since $ax^2 + by^2 + cz^2 = 0$, $p \mid ax^2$. Since $p \mid a$ and $\gcd(a, c) = 1$, $p \mid x^2$ and thus $p \mid x$, contradicting $\gcd(x, y, z) = 1$. So we may assume that at least one of y and z is invertible modulo p; with no real loss of generality we assume y is invertible modulo p. Then $by^2 \equiv -cz^2 \pmod{p}$, so

$$-bc \equiv \left(\frac{cz}{y}\right)^2 \pmod{p},$$

i.e., $-bc$ is a square modulo p. Since this argument holds for every prime dividing the squarefree integer a, by the Chinese Remainder Theorem $-bc$ is a square modulo a. And of course a perfectly symmetrical argument shows that $-ac$ is a square modulo b and that $-ab$ is a square modulo c.

Remarkably, these easy necessary conditions and also sufficient.

Lemma 16.2. *Let $m \in \mathbb{Z}^+$ and let $\epsilon_1, \epsilon_2, \epsilon_3 \in \mathbb{R}^{>0}$ be such that $\epsilon_1\epsilon_2\epsilon_3 \geq m$. Let $\ell(x, y, z) = \alpha x + \beta y + \gamma z \in \mathbb{Z}[x, y, z]$ be any linear polynomial. Then there are $(x, y, z) \in (\mathbb{Z}^3)^\bullet$ such that*

$$\ell(x, y, z) \equiv 0 \pmod{m} \tag{53}$$

and $|x| \leq \epsilon_1$, $|y| \leq \epsilon_2$, $|z| \leq \epsilon_3$.

Exercise: Prove Lemma 16.2. (Suggestion: show that (53) defines a sublattice $\Lambda \subset \mathbb{Z}^3$ of index dividing m, and apply Minkowski's Linear Forms Theorem.)

Theorem 16.3. *(Legendre) The Legendre Equation has a nontrivial integer solution iff $-bc$ is a square mod a, $-ac$ is a square mod $|b|$ and $-ab$ is a square mod $|c|$.*

Proof. We may assume that b and c are not both -1. Indeed, if $b = c = -1$, then the condition $-bc$ is a square modulo a gives that -1 is a square modulo a and thus a is a sum of two integer squares, yielding a nontrivial solution to $ax^2 - y^2 - z^2 = 0$.

We claim that our congruence conditions force $q(x, y, z)$ are necessary and sufficient for the existence of linear forms $L_1(x, y, z), L_2(x, y, z) \in \mathbb{Z}[x, y, z]$ such that

$$q(x, y, z) \equiv L_1(x, y, z)L_2(x, y, z) \pmod{abc}.$$

Since a, b, c are coprime in pairs, it is sufficient to show the factorization of q into linear forms modulo a, modulo b and modulo c; then by the Chinese Remainder Theorem we may choose $L_1, L_2 \in \mathbb{Z}[x, y, z]$ which reduce modulo a, b and c to the linear factors of q. So: let r be such that $r^2 \equiv -bc \pmod{a}$, and let c' be such that $cc' \equiv 1 \pmod{a}$. Then[16]

$$q(x, y, z) = ax^2 + by^2 + cz^2 \equiv by^2 + cz^2 \equiv cc'(by^2 + cz^2) \equiv c'(bcy^2 + c^2z^2)$$

[16]That these congruence conditions imply factorizations of binary forms is not a new phenonemon for us: it was established in some generality in Proposition 9.10. Unfortunately for us, the generality of that result is not the "right generality" for the current application: Proposition 9.10 concerns quadratic forms over domains of characteristic not 2, whereas we are looking at quadratic forms over rings $\mathbb{Z}/(a)$, which may have zero-divisors...including 2.

$$\equiv c'(c^2z^2 - r^2y^2) \equiv c'(cz + ry)(cz - ry) \equiv L_1(x,y,z)L_2(x,y,z) \pmod{a}.$$

By symmetry similar arguments can be made modulo b and c. So we get

$$q(x,y,z) \equiv L_1(x,y,z)L_2(x,y,z) = (\alpha x + \beta y + \gamma z)(\alpha' x + \beta' y + \gamma' z) \pmod{abc}.$$

Now apply Lemma 16.2 with $m = abc$, $\epsilon_1 = \sqrt{|bc|}$, $\epsilon_2 = \sqrt{|ac|}$, $\epsilon_3 = \sqrt{|bc|}$: there are $(x_1, y_1, z_1) \in (\mathbb{Z}^3)^\bullet$ with

$$|x_1| \leq \sqrt{bc},\ |x_2| \leq \sqrt{ac},\ |x_3| \leq \sqrt{ab} \tag{54}$$

and

$$L_1(x_1, y_1, z_1) \equiv 0 \pmod{abc}.$$

Since $q \equiv L_1 L_2 \pmod{abc}$, this implies

$$q(x_1, y_1, z_1) \equiv 0 \pmod{abc}.$$

Note that we have

$$x_1^2 \leq bc\ y_1^2 \leq -ac,\ z_1^2 \leq -ab. \tag{55}$$

In fact, since bc is squarefree and greater than 1, we must have $x_1^2 < bc$. Similarly, if $y_1^2 = -ac$ then $a = 1$ and $c = -1$, and if $z_1^2 = -ab$ then $a = 1$ and $b = -1$, so at least one of the two inequalities must be strict and thus

$$-2abc < by_1^2 + cz_1^2 \leq ax_1^2 + by_1^2 + cz_1^2 \leq ax_1^2 < abc.$$

Thus either $q(x_1, y_1, z_1) = 0$ – great! – or $q(x_1, y_1, z_1) = -abc$. In the latter case, the ternary form q represents $-\operatorname{disc}(q)$ hence is isotropic by [NCA, Cor. 95].[17] If one wants to avoid this nontrivial result of quadratic form theory, here is a completely elementary finish: put

$$x_2 = -by_1 + x_1 z_1,$$
$$y_2 = ax_1 + y_1 z_1,$$
$$z_2 = z_1^2 + ab.$$

Then

$$q(x_2, y_2, z_2) = ab(ax_1^2 + by_1^2 + cz_1^2) + z_1^2(ax_1^2 + by_1^2 + cz_1^2) + abcz_1^2 + a^2b^2c$$
$$= ab(-abc) - abcz_1^2 + abcz_1^2 + a^2b^2c = 0,$$

so (x_2, y_2, z_2) is a solution. If $z_2 = z_1^2 + ab = 0$ then $a = 1$, $b = -1$, and $(1, 1, 0)$ is a nontrivial solution. □

Now let us make some remarks and further inquiries about this proof.

First, the argument begins in a similar way to our study of numbers *represented* by binary and quaternary forms, but instead of choosing a (suitable) number d and constructing a sublattice of index some power of d, here we are constructing once and for all a sublattice Λ_{abc} of index abc based upon properties of q modulo **singular primes**, i.e., primes dividing $\operatorname{disc} q = abc$.

Further, the proof is cast in terms of factorization of q into linear forms, but it is possible to recast it in the (by now) more familiar language of magic lattices. Namely, we can more directly construct the magic lattice Λ of index abc as follows: a Chinese Remainder Theorem argument reduces us to constructing a lattice Λ_p for all $p \mid abc$ of index p such that $q|_{\Lambda_p} \equiv 0 \pmod{p}$. When p is an *odd* prime, this

[17] I am indebted to Danny Krashen for pointing out this simplification of the end of the proof.

can be established easily as follows: by hypothesis, p divides exactly one of a, b, c, so without loss of generality suppose $p \mid c$; then the reduction of q modulo p is the *ternary form* $ax^2 + by^2 + 0z^2$. Thus $q = q' \oplus R(q)$, the orthogonal direct sum of a nondegenerate binary quadratic form $q' = ax^2 + by^2$ and a one-dimensional isotropic subspace $R(q)$. Under the reduction map $\mathbb{Z}^3 \to (\mathbb{Z}/p\mathbb{Z})^3$, the index p sublattices correspond to the codimension one – hence dimension 2 – subspaces of $(\mathbb{Z}/p\mathbb{Z})^3$, and we are looking for a codimension one subspace on which q is identically zero (modulo p). Of course $R(q)$ gives a one-dimensional isotropic subspace, so we need one more dimension. We can get this iff $q' = ax^2 + by^2$ is itself isotropic, iff its discriminant is minus a square modulo p. But we have assumed that $-ab$ is a square modulo c, so in particular it is a square modulo p. This shows that not only does the required two-dimensional subspace exist, it is *unique*.

We need to make a separate argument when $p = 2$, since quadratic form theory works quite differently in charateristic 2. Fortunately this is a very simple situation: we may assume a and b are odd and c is even, so modulo 2 we have $q(x, y, z) = x^2 + y^2 + 0z^2$, and we can see right away that q vanishes identically on the subspace $\{0, (1, 1, 0), (0, 0, 1), (1, 1, 1)\}$.[18]

The idea of applying Minkowski's Linear Forms Theorem is a clever one. From our perspective it would be more natural to apply the Convex Body Theorem. But we have a homogeneous indefinite quadratic equation – so where is our convex body?? Here is one simple, classical idea: **majorization**. Namely, let us consider the quadratic form $|q|(x, y, z) = |a|x^2 + |b|y^2 + |c|z^2$. Evidently this form is positive definite and bears some relation to the indefinite form q: more precisely it *majorizes* q in the sense that for all $(x, y, z) \in \mathbb{R}^3$,

$$|q(x, y, z)| \leq |q|(x, y, z).$$

We are therefore free to apply the Convex Body Theorem to the level sets $\Omega_R = \{v \in \mathbb{R}^3 \mid |q|(v) \leq R^2\}$ and the lattice Λ_{abc}. If there exists $R^2 < abc$ and $v \in \Omega_R \cap \Lambda_{abc} \neq \{0\}$, $|q(v)| \leq |q|(v) < abc$ and $q(v) \equiv 0 \pmod{abc}$, so $q(v) = 0$.

Unfortunately it does not quite work: we have $\operatorname{disc}|q| = \operatorname{Covol}\Lambda_{abc} = abc$, so applying Theorem 9.6 we get $v \in \Lambda_{abc}^{\bullet}$ with

$$|q|(v) \leq \frac{4(\operatorname{disc}|q|)^{\frac{1}{3}}}{V_3^{\frac{2}{3}}} (\operatorname{Covol}\Lambda_{abc})^{\frac{2}{3}} = \frac{4}{(4\pi/3)^{2/3}} abc = (1.5393\ldots)abc.$$

Too bad – we needed the coefficient of abc to be less than one! Note that even if we replaced the Minkowski constant M_3 by the 3-dimensional Hermite constant $\gamma_3 = 2^{\frac{1}{3}}$, we still don't quite get what we want:

$$|q|(v) \leq \gamma_3 abc = (1.25992\ldots)abc.$$

Thus it seems that the trick of switching from (x_1, y_1, z_1) to the pulled-out-of-thin-air (x_2, y_2, z_2) is necessary to complete the argument. This is disappointing, because we had a beautiful upper bound on the size of (x_1, y_1, z_1) with respect to a weighted ℓ_∞-norm on $\mathbb{R}^3$, which passage to (x_2, y_2, z_2) ruins completely (we can still get an explicit upper bound, but a much worse one).

[18]In fact, over $\mathbb{F}_2$ we have $x^2 = x$, so q is "really" the nonzero linear form $x + y$, so without any calculation it is clear that it has a codimension one kernel.

16.2.1. *Transcription of* [DH48].

Let a, b, c be integers, and suppose that (i) not all of a, b, c have the same sign, and none of them is zero, (ii) a, b, c are relatively prime in pairs. A famous theorem of Legendre asserts that if the congruences

$$A^2 \equiv -bc \pmod{a},\ B^2 \equiv -ac \pmod{b},\ C^2 \equiv -ab \pmod{c} \tag{56}$$

are all soluble, then the equation

$$ax^2 + by^2 + cz^2 = 0 \tag{57}$$

has a solution in integers x, y, z, not all zero. If the additional hypothesis is made that a, b, c are squarefree, then the above condition is necessary as well as sufficient, for the solubility of (57). In this note we give a simple proof of Legendre's theorem by using the methods of the geometry of numbers. The idea of the proof is the natural one of determining x, y, z so that $ax^2 + by^2 + cz^2$ is both divisible by abc and numerically less than $|abc|$.

The points (x, y, z), where x, y, z are any integers satisfying

$$Ay \equiv cz \pmod{a},\ Bz \equiv ax \pmod{c},\ Cx \equiv by \pmod{c}, \tag{58}$$

form a lattice in three-dimensional space, since the sum or difference of two such points is again such a point. Three particular lattice points are

$$(bc, 0, 0),\ (x_1, a, 0),\ (x_2, y_2, 1)$$

for appropriate values of x_1, x_2, y_2. Moreover every solution of (58) is expressible in the form

$$(x, y, z) = \lambda(bc, 0, 0) + \mu(x_1, a_1, 0) + \nu(x_2, y_2, 1)$$

with integral λ, μ, ν, as one easily sees by determining ν, μ, λ successively. Since the determinant of the coordinates of the three points is abc, it follows that the determinant of the lattice is $|abc|$. All points of the lattice satisfy

$$ax^2 + by^2 + cz^2 \equiv 0 \pmod{abc}. \tag{59}$$

For

$$c(b^2 + cz^2) \equiv -A^2y^2 + c^2z^2 \equiv 0 \pmod{a}$$

by (56) and (58), whence (59) follows as a congruence to modulus a and similarly to moduli b and c.

supppose, as we may after (i), that $a > 0,\ b > 0,\ c = -c' < 0$. The real linear transformation

$$x = X\sqrt{bc'},\ y = Y\sqrt{ac'},\ z = Z\sqrt{ab} \tag{60}$$

transforms the lattice in (x, y, z)-space into a lattice in (X, Y, Z)-space whose determinant is 1. For every point of this new lattice,

$$X^2 + Y^2 - Z^2 = (ax^2 + by^2 + cz^2)/(abc'),$$

and so is an integer. To complete the proof of Legendre's theorem, it will suffice to prove that every lattice of determinant 1 contains a point, other than the origin 0,

which satisfies[19]

$$|X^2 + Y^2 - Z^2| < 1. \tag{61}$$

Let M denote the lower bound of $|X^2 + Y^2 - Z^2|$ for all points of the lattice other than 0. If $M < 1$, the desired conclusion holds, so we may suppose that $M \geq 1$. There exists a lattice point (X_1, Y_1, Z_1) for which

$$X_1^2 + Y_1^2 - Z_1^2 = \pm M_1, \tag{62}$$

where M_1 either equals M,, or differs from m by an arbitrarily small amount.

(Comment by PLC: They seem to have forgotten that $X^2 + Y^2 - Z^2$ is integer-valued on the lattice. We can thus take $M_1 = M$.)

Case 1. Suppose first that the lower sign holds in (62). We can find a real inear transformation which will leave $X^2+Y^2-Z^2$ invariant, and transform (X_1, Y_1, Z_1) into $(0, 0, \sqrt{M_1})$. For we can first transform (X_1, Y_1, Z_1) into $(0, \sqrt{X_1^2 + Y_1^2}, Z_1)$ by a trnasofrmation of X, Y only, leaving $X^2 + Y^2$ invariant, and then transform $(0, \sqrt{X_1^2 + Y_1^2}, Z_1)$ into $(0, 0, \sqrt{M_1})$ by a transformation of Y, Z only, leaving $Y^2 - Z^2$ invariant.

(Comment by PLC: alternately, $X^2 + Y^2 - Z^2$ is a nondegenerate quadratic form, so by Witt's isometry extension theorem, the orthogonal group of this form acts transitively on the set of all vectors taking a given nonzero value.)

We now have a lattice in (X, Y, Z)-space, of determinant 1, such that $(0, 0, \sqrt{M_1})$ is a lattice point, and such that

$$|X^2 + Y^2 - Z^2| \geq M \tag{63}$$

for every lattice point except 0. The points $(X, Y, 0)$ obtained by projecting all lattice points on the plane $Z = 0$ form a two-dimensional lattice of determinant $1/\sqrt{M_1}$. We apply the well-known theorem that any plane lattice of determinant Δ contains a lattice point, other than O, in the circle $X^2 + Y^2 \leq 2\Delta/\sqrt{3}$. Thus there is a point of the three-dimensional lattice satisfying

$$X^2 + Y^2 \leq 2/\sqrt{3M_1}. \tag{64}$$

We can suppose, without loss of generality, by subtracting a multiple of $(0, 0, \sqrt{M_1})$ and changing signs throughout if necessary, that

$$0 \leq Z \leq \frac{1}{2}\sqrt{M_1}. \tag{65}$$

Thus $X^2 + Y^2 - Z^2 \geq -\frac{1}{4}M_1$, and since M_1, is approximately M, (63)implies that

$$X^2 + Y^2 - Z^2 \geq M. \tag{66}$$

Another lattice point is $(X, Y, Z - \sqrt{M_1})$, and by (64) and (65),

$$X^2 + Y^2 - (Z - \sqrt{M_1})^2 \leq 2/\sqrt{3M_1} - \frac{1}{4}M_1.$$

[19]In fact, by a theorem of Markoff, this is true with any number greater than $\sqrt{2/3}$ in place of 1 on the right of 61). For an elementary proof of Markoff's theorem, see Davenport, *J. of London Math. Soc.* 22 (1947), 96-9.

Since M_1 is approximately M, and $M \geq 1$, the number on the right is less than M. Hence (63) gives

$$X^2 + Y^2 - (Z - \sqrt{M_1})^2 \leq -M. \tag{67}$$

On subtraction, (66) and (67) imply

$$2Z\sqrt{m_1} \leq M_1 - 2M \leq 0,$$

contrary to (65).

Case 2. Supose now that the upper sign holds in (62). We can suppose also that there are no values of $X^2 + Y^2 - Z^2$ arbitrarily near to $-M$, since that possibility has been settled in Case 1. We can find (in the same way as befoe) a real linear transformation which leaves $X^2 + Y^2 - Z^2$ invariant and transforms (X_1, Y_1, Z_1) into $(\sqrt{M_1}, 0, 0)$. Thus we have a lattice in (X, Y, Z)-space of determinant 1 such that $(\sqrt{M_1}, 0, 0)$ is a lattice point, and such that (63) holds for all lattice points except 0. The points $(0, Y, Z)$ obtained by projecting the lattice points on the plane $X = 0$, form a lattice of determinant $1/\sqrt{M_1}$. We use the well-known theorem that a plane lattice of determinant Δ has a point, other than O, in the rectangle

$$|Y| < \lambda,\ |Z| < \mu$$

if λ, μ are positive numbers satisfying $\lambda\mu > \Delta$. Hence there is a point of the three-dimensional lattice satisfying

$$|Y| < (M - \frac{1}{4}M_1)^{\frac{1}{2}},\ |Z| < M + M_1.$$

For, since M_1 is approximately M and $M \geq 1$, the product of the numbers on the right is greater than $1/\sqrt{M_1}$. We can suppose without loss of generality that this lattice point has

$$0 \leq X \leq \frac{1}{2}\sqrt{M_1}. \tag{68}$$

Now $X^2 + Y^2 - Z^2 < \frac{1}{4}M_1 + (M - \frac{1}{4}M_1) = M$; hence (63) implies

$$X^2 + Y^2 - Z^2 \leq -M. \tag{69}$$

Another lattice point is $(X + \sqrt{M_1}, Y, Z)$ and

$$(X + \sqrt{M_1})^2 + Y^2 - Z^2 > M_1 - (M + M_1) = -M.$$

Hence (63) gives

$$(X + \sqrt{M_1})^2 + Y^2 - Z^2 \geq M. \tag{70}$$

By (69) and (70), $2X\sqrt{M_1} \geq 2M - M_1$. So, by (68), X is nearly $\frac{1}{2}\sqrt{M}$. Also

$$M - (X + \sqrt{M_1})^2 \leq Y^2 - Z^2 \leq -M - X^2.$$

So $Y^2 - Z^2$ is nearly $-\frac{5}{4}M$. But then $X^2 + Y^2 - Z^2$ is nearly $-M$, which contradicts the hypothesis made earlier.

16.2.2. *Notes on* [Mo51].

Several authors have improved the endgame of the proof of Legendre's Theorem so as to recover not the full bound (55) but something which is weaker only by a constant factor. Of these, the one which we find the most interesting and thematic is due to Mordell [Mo51], and we present Mordell's argument in detail here.[20]

Now we give Mordell's modification of the GoN proof of Legendre's Theorem. We will follow Mordell's setup, which is to write the Legendre Equation as

$$f(x,y,z) = ax^2 + by^2 - cz^2 = 0$$

with $a, b, c \in \mathbb{Z}^+$, abc squarefree.

Theorem 16.4. *(Mordell* [Mo51]*) Assume the Legendre Conditions. There is* $(x_0, y_0, z_0) \in (\mathbb{Z}^3)^\bullet$ *with*

$$f(x_0, y_0, z_0) \equiv 0 \pmod{4abc}$$

and

$$|x_0| \leq \sqrt{2bc},\ |y_0| < \sqrt{2ac},\ |z_0| < 2\sqrt{ab}.$$

Before giving the proof, we remark that Theorem 16.4 implies Legendre's Theorem: since $f(x_0, y_0, z_0) \equiv 0 \pmod{4abc}$ and

$$|f(x_0, y_0, z_0)| = |ax_0^2 + by_0^2 - cz_0^2| \leq \max(ax_0^2 + by_0^2, cz_0^2) < 4abc,$$

we must have $f(x_0, y_0, z_0) = 0$. Further – and this is really the point – we also get a bound on the coefficients of an isotropic vector that is within a factor of 2 of the bound (54) obtained – and then lost! – in our first proof of Legendre's Theorem.

Proof. The basic strategy of the proof is to use Theorem 12.7 with $m = 5$, $n = 3$, $\epsilon_1 = \sqrt{2bc}$, $\epsilon_2 = \sqrt{2ac}$, $\epsilon_3 = 2\sqrt{ab}$. We write out the Legendre Conditions: there are integers A, B, C such that

$$bA^2 \equiv c \pmod{a},\ cB^2 \equiv a \pmod{b},\ aC^2 \equiv -b \pmod{c}. \tag{71}$$

Case I: Suppose abc is odd. As the first three out of five congruences, we take

$$y - Az \equiv 0 \pmod{a},\ z - Bx \equiv 0 \pmod{b},\ x - Cy \equiv 0 \pmod{c}. \tag{72}$$

Then

$$by^2 - cz^2 \equiv b(y^2 - A^2z^2) \equiv 0 \pmod{a},$$
$$ax^2 + by^2 \equiv a(x^2 - C^2y^2) \equiv 0 \pmod{c},$$
$$ax^2 - cz^2 \equiv c(B^2x^2 - z^2) \equiv 0 \pmod{b},$$

so that all $(x, y, z) \in \mathbb{Z}^3$ satisfying (72) also satisfy

$$f(x,y,z) \equiv 0 \pmod{abc}. \tag{73}$$

(This part of the argument is, of course, familiar to us: it is equivalent to the construction of the lattice Λ_{abc}.) Since abc is odd, to get $f(x, y, z) \equiv 0 \pmod{4abc}$ it suffices to impose congruence conditions which imply $f(x, y, z) \equiv 0 \pmod{4}$, and this will be done by quite elementary means.

[20] In [Mo69], Mordell comments that essentially the same approach as his was taken independently by Skolem in the slightly later paper [Sk52].

First we CLAIM that we *cannot* have $a \equiv b \equiv -c \pmod 4$, as we show by the following Jacobi symbol calculation: using (71) we find

$$\left(\frac{-ab}{c}\right) = \left(\frac{(-b)(-b)}{c}\right) = 1,$$

$$\left(\frac{bc}{a}\right) = \left(\frac{b^2}{a}\right) = 1,$$

$$\left(\frac{ac}{b}\right) = \left(\frac{c^2}{b}\right) = 1,$$

so

$$1 = \left(\frac{-1}{c}\right)\left(\left(\frac{a}{b}\right)\left(\frac{b}{a}\right)\right)\left(\left(\frac{b}{c}\right)\left(\frac{c}{b}\right)\right)\left(\left(\frac{c}{a}\right)\left(\frac{a}{c}\right)\right).$$

• If $c \equiv 3 \pmod 4$ – so $a \equiv b \equiv 1 \pmod 4$ – then $\left(\frac{-1}{c}\right) = -1$ and by Quadratic Reciprocity for the Jacobi symbol the rest of the product evaluates to 1, giving $1 = -1$, a contradiction.
• If $c \equiv 1 \pmod 4$ – so $a \equiv b \equiv 3 \pmod 4$ – then

$$\left(\frac{-1}{c}\right) = \left(\frac{b}{c}\right)\left(\frac{b}{c}\right) = \left(\frac{c}{a}\right)\left(\frac{a}{c}\right) = 1,$$

while

$$\left(\frac{a}{b}\right)\left(\frac{b}{a}\right) = -1,$$

contradiction. Thus two of $a, b, -c$ are incongruent modulo 4.
• Suppose $a \not\equiv b \pmod 4$. Then $a \equiv -b \pmod 4$ and we take as our last two congruences

$$x \equiv y \pmod 2,\ z \equiv 0 \pmod 2,$$

so that $f(x,y,z) \equiv ax^2 + by^2 - cz^2 \equiv a(x^2 - y^2) \equiv 0 \pmod 4$.
• Suppose $a \equiv -c \pmod 4$. Then $a \equiv c \pmod 4$ and we take as our last two congruences

$$x \equiv z \pmod 2,\ y \equiv 0 \pmod 2.$$

• Suppose $b \equiv -c \pmod 4$. Then $b \equiv c \pmod 4$ and we take as our last two congruences

$$y \equiv z \pmod 2,\ x \equiv 0 \pmod 2.$$

In all three cases we apply Theorem 12.7 with $d_1 = a$, $d_2 = b$, $d_3 = c$, $d_4 = d_5 = 2$, to get the desired result.
Case II: Suppose $2 \mid a$. We take now as our first three congruences

$$y - Az \equiv 0 \pmod{\frac{a}{2}},\ z - Bx \equiv 0 \pmod b,\ x - Cy \equiv 0 \pmod c, \tag{74}$$

which as above imply

$$f(x,y,z) \equiv 0 \pmod{\frac{abc}{2}}. \tag{75}$$

We now wish to impose congruences modulo 2 and modulo 4 which imply $f(x,y,z) \equiv 0 \pmod 8$ and apply Theorem 12.7 with $d_1 = \frac{a}{2}$, $d_2 = b$, $d_3 = c$, $d_4 = 4$, $d_5 = 2$. First we observe, as above, that

$$\left(\frac{bc}{a/2}\right) = \left(\frac{b^2}{a/2}\right) = 1,\ \left(\frac{ac}{b}\right) = 1,\ \left(\frac{-ab}{c}\right) = 1,$$

so

$$1 = \left(\frac{-1}{c}\right)\left(\frac{2}{bc}\right)\left(\left(\frac{a/2}{bc}\right)\left(\frac{bc}{a/2}\right)\right)\left(\left(\frac{b}{c}\right)\left(\frac{c}{b}\right)\right). \tag{76}$$

Case 1: Suppose $a + b - c \equiv 0 \pmod 8$. Then we impose

$$x \equiv y \pmod 2,\ z \equiv y \pmod 4,$$

and

$$f(x,y,z) = ax^2 + by^2 - cz^2 \equiv ax^2 + by^2 - (a+b)y^2 \equiv a(x^2 - y^2) \equiv 0 \pmod 8.$$

Case 2: Suppose $b - c \equiv 0 \pmod 8$. Then we impose

$$x \equiv 0 \pmod 2,\ z \equiv y \pmod 4,$$

and

$$f(x,y,z) = ax^2 + by^2 - cz^2 \equiv ax^2 + by^2 - cy^2 \equiv ax^2 \equiv 0 \pmod 8.$$

Case 3: Suppose $b - c \equiv 4 \pmod 8$, so $bc \equiv \pm 3 \pmod 8$.
• If $b \equiv c \equiv 1 \pmod 4$, then all factors in (76) are 1, except $\left(\frac{2}{bc}\right) = 1$, contradiction.
• If $b \equiv c \equiv 3 \pmod 4$, all factors in (76) are -1 except $\left(\left(\frac{a/2}{bc}\right)\left(\frac{bc}{a/2}\right)\right) = 1$, contradiction.
Case 4: Suppose $a \equiv b - c \pmod 8$, so $b \equiv \pm 1 \pmod 4$ and $c \equiv \mp 1 \pmod 4$.
Case 4a): Suppose $a \equiv 2 \pmod 8$. Then $\left(\left(\frac{a/2}{bc}\right)\left(\frac{bc}{a/2}\right)\right) = \left(\left(\frac{b}{c}\right)\left(\frac{c}{b}\right)\right) = 1$, whereas an examination of the four possible cases

$$(b,c) \equiv (3,1),\ (5,3),\ (7,5),\ (1,7) \pmod 8$$

shows that we always have $\left(\frac{-1}{c}\right)\left(\frac{2}{bc}\right) = -1$, contradiction.
Case 4b): Suppose $a \equiv 6 \pmod 8$. A similar examination of all four possible cases

$$(b,c) \equiv (3,5),\ (5,7),\ (7,1),\ (1,3) \pmod 8$$

leads to a contradiction.
Case III: Suppose $2 \mid b$. This case follows from the $2 \mid a$ case by symmetry.
Case IV: Suppose $2 \mid c$. We leave this case as an exercise, as did Mordell. □

Exercise: Fill in the details of Cases II4b) and IV.

Exercise: a) Show that Mordell's proof yields the existence of a sublattice $\Lambda' \subset \mathbb{Z}^3$, of index $4abc$, such that $q \equiv 0 \pmod{4abc}$ on Λ'.
b) Use part a) and Theorem 9.6 to deduce Legendre's Theorem.

16.3. **Holzer's Theorem.**

Theorem 16.5. *(Holzer* [Ho50]*) Let $a, b, c \in \mathbb{Z}^\bullet$ be such that abc is squarefree. If the Legendre form $q = ax^2 + by^2 + cz^2 = 0$ admits an isotropic vector, it admits an isotropic vector $v = (x, y, z)$ with*

$$|x| \leq \sqrt{|bc|},\ |y| \leq \sqrt{|ac|},\ |z| \leq \sqrt{|ab|}.$$

Remark: Holzer's bound implies the existence of an isotropic vector v with

$$|v| = \max |x|, |y|, |z| \leq \max |a|, |b|, |c|.$$

It is natural to compare this to the bound given by the Cassels Isotropy Theorem: applying that result to q, we get an isotropic vector $v = (x, y, z)$ with

$$|v| \leq 3(|a| + |b| + |c|).$$

Thus when both apply, Holzer's bound is an *improvement* of Cassels's bound.

Holzer's own proof of Theorem 16.5 used rather advanced algebraic number theory, including results of Hecke generalizing Dirichlet's Theorem on primes in arithmetic progressions to the case of number fields.

For some years after Holzer's work various well known mathematicians gave elementary proofs of bounds slightly weaker than Holzer's bound: see e.g. [Mo51], [Sk52], [C, p. 102]. Finally, in one of his last works, Mordell gave in 1969 a reasonably short, completely elementary (but rather tricky) proof of Holzer's bound [Mo69]. Mordell's argument has been found confusing by some, but a nice exposition of it was given by L.M. Nunley [Nu10, §2.4].

Note well that the Holzer bound is precisely the bound we were getting in the GoN proof of Legendre's Theorem, until we lost it at the very end by making a mysterious change of variables. This suggests the goal of giving an alternate endgame to this proof which yields a proof of Legendre's Theorem and Holzer's Theorem in one fell swoop. Exactly this was done by Cochrane and Mitchell: we present their argument in the next section. Here we content ourself with further remarks, examples and strengthenings.

Example ([Nu10, p. 14]): Let p be a prime, and consider the Legendre equation

$$x^2 + y^2 - pz^2 = 0.$$

The Legendre Conditions hold iff -1 is a square modulo p iff $p = 2$ or $p \equiv 1 \pmod 4$. (From this we deduce that all such primes are sums of two *rational* squares. Since $x^2 + y^2$ is a Euclidean form, it is an ADC form, and thus we recover the Two Squares Theorem.) Here the Holzer bound is

$$|x| \leq p, |y| \leq p, |z| \leq 1.$$

This shows that the Holzer bound cannot be improved to

$$|x| \leq C\sqrt{|bc|},\ |y| \leq C\sqrt{|ac|},\ |z| \leq C\sqrt{|ab|}$$

for any constant $C < 1$: indeed, such an improvement would in the above examples force $z = 0$ and thus $x = y = 0$.

There are also infinitely many cases where the Holzer bound is not sharp.

Theorem 16.6. *(*[Nu10, p. 17]*) For any $C > 0$, there are infinitely many Legendre equations $ax^2 + by^2 + cz^2 = 0$ admitting an isotropic vector $v = (x, y, z)$ with*

$$|x| \leq C\sqrt{|bc|},\ |y| \leq C\sqrt{|ac|},\ |z| \leq C\sqrt{|ab|}.$$

To the best of my knowledge, the only published work which seriously explores going beyond the Holzer bound under suitable conditions is a 1959 paper of Kneser [Kn59]. Kneser's paper has no mathscinet citations: I learned of it through [Wi88]. I have not yet gotten hands on Kneser's paper, and to add injury to that insult it is

written in German, so it would be difficult for me to understand. Perhaps someone would like to tell me what's in it?

Corollary 16.7. *Let $a, b, c \in \mathbb{Z}^\bullet$, and put $d = \gcd(a, b, c)$. Suppose*

$$q = ax^2 + by^2 + cz^2 = 0$$

is an isotropic form. Then q has an isotropic vector $v = (x, y, z)$ satisfying

$$|x| \leq \frac{\sqrt{|bc|}}{d},\ |y| \leq \frac{\sqrt{|ac|}}{d},\ |z| \leq \frac{\sqrt{|ab|}}{d}. \tag{77}$$

Proof. We go by strong induction on $N = |abc|$.
Step 0 (Base Case): When $N = 1$, then $|a| = |b| = |c| = 1$; this is a trivial case.

Henceforth we assume $N > 1$ and, inductively, that the result holds for all isotropic forms $a'x^2 + b'y^2 + c'z^2 = 0$ with $|a'b'c'| < N$.
Step 1: Suppose $d = \gcd(a, b, c) > 1$. Let $q' = \frac{a}{d}x^2 + \frac{b}{d}y^2 + \frac{c}{d}z^2$. Since $q = dq'$, q and q' have the same isotropic vectors. In particular, since q is isotropic, so is q', and by induction there is an isotropic vector $v = (x, y, z)$ for q' with

$$|x| \leq \sqrt{\left|\frac{b}{d}\right|\left|\frac{c}{d}\right|} = \frac{\sqrt{|bc|}}{d},$$

$$|y| \leq \sqrt{\left|\frac{a}{d}\right|\left|\frac{c}{d}\right|} = \frac{\sqrt{|ac|}}{d},$$

$$|z| \leq \sqrt{\left|\frac{a}{d}\right|\left|\frac{b}{d}\right|} = \frac{\sqrt{|ab|}}{d}.$$

Since v is also an isotropic vector for q, we're done in this case.
Step 2: Suppose $d = 1$ and that a, b and c are not all squarefree; without loss of generality, suppose $a = p^2a'$ for some prime number p. Then

$$0 = q(v) = p^2a'x^2 + by^2 + cz^2 = a'(px)^2 + by^2 + cz^2,$$

so the form $q'(x, y, z) = a'x^2 + by^2 + cz^2$ is isotropic. Since $|a'bc| = \frac{N}{p^2} < 1$, by induction there are $x_0, y_0, z_0 \in \mathbb{Z}$, not all 0, such that $a'x_0^2 + by_0^2 + cz_0^2 = 0$ and

$$|x_0| \leq \sqrt{|bc|},\ |y_0| \leq \sqrt{|a'c|} = \frac{\sqrt{|ac|}}{p},\ |z_0| \leq \sqrt{|a'b|} = \frac{\sqrt{|ab|}}{p}.$$

Take $v = (x_0, py_0, pz_0)$ is an isotropic vector for q satisfying (77).
Step 3: Suppose $d = 1$, a, b, c are all squarefree, but that they are *not* pairwise coprime; without loss of generality, $a = pa'$, $b = pb'$ for some prime number p. If $v = (x, y, z)$ is an isotropic vector for q then $p \mid z$. Thus we may take $z = pz'$. Since

$$q(v) = pa'x^2 + pb'y^2 + p^2cz'^2 = 0,$$

we find

$$0 = a'x^2 + b'y^2 + pcz'^2 = q'(v'),$$

where $q'(x, y, z) = a'x^2 + b'y^2 + c'z^2$, $c' = pc$, and $v' = (x, y, \frac{z}{p})$. Since $|a'b'c'| = |\frac{abc}{p}| < N$, by induction there are $x_0, y_0, z_0 \in \mathbb{Z}$, not all zero, such that

$$\frac{a}{p}x_0^2 + \frac{b}{p}y_0^2 + cpz_0^2 = 0,$$

$$|x_0| \leq \sqrt{|bc|},\ |y_0| \leq \sqrt{|ac|}, |z_0| \leq \sqrt{|ab|}p.$$

Then $v = (x_0, y_0, pz_0)$ is an anisotropic vector for q satisfying (77).
Step 4: Suppose a, b, c are squarefree and pairwise coprime. Then q is in Legendre Form and Theorem 16.9 applies. □

Corollary 16.7 appears in [Wi88]: therein, Williams modifies the proof of [Mo69] to go through without the reduction to Legendre form. (For those who are tracking down the reference, the bound given there is superficially different, but equivalent via Proposition 16.8 of the next section.) However, as Cochrane and Mitchell explain in [CoMi98], one may simply deduce the result from Holzer's Theorem by tracking through the isotropic vector through the reduction process. This is of course the strategy followed in the above proof (with a slightly different implementation).

In view of Corollary 16.7 one may wonder why we bother to reduce diagonal ternary quadratic forms into Legendre form. There is a good answer to this: for arbitrary nonzero integers $a, b, c \in \mathbb{Z}^\bullet$, not all of the same sign, the Legendre conditions

- $-ab$ is a square modulo c,
- $-ac$ is a square modulo b,
- $-bc$ is a square modulo a

are *insufficient* to force isotropy. For example, let p be a prime number and consider

$$px^2 + py^2 - z^2 = 0.$$

In this case the Legendre conditions are satisfied for any prime p. However, the corresponding Legendre form is

$$x^2 + y^2 - pz^2 = 0,$$

which is anisotropic if $p \equiv 3 \pmod 4$: indeed, applying the Legendre Conditions to the Legendre form we get that -1 needs to be a square modulo p.

16.4. **The Cochrane-Mitchell Theorem.**

The above goal of giving a simultaneous proof of Legendre's Theorem and the Holzer bound was attained in a beautiful relatively recent work of Cochrane and Mitchell [CoMi98]. Their key idea is to refine the lattice Λ to a suitable sublattice Λ'. Suppose for the sake of argument that we can find, for some $n > 1$, an index n sublattice Λ' of Λ with $q|_{\Lambda'} \equiv 0 \pmod{abcn}$. Then there exists $v \in \Lambda'^\bullet$ with

$$|q(v)| \leq |q|(v) \leq 2^{\frac{1}{3}} (\operatorname{disc} |q|)^{\frac{1}{3}} (\operatorname{Covol} \Lambda')^{\frac{2}{3}} = 2^{\frac{1}{3}} n^{\frac{2}{3}} abc.$$

When $n = 2$ we have $|q(v)| \leq 2abc$, and is this right on the boundary: if we can also show that $q|_{\Lambda'}$ is not H-equivalent to $q_3 = x^2 + y^2 + z^2 + xy + xz + yz$, then Gauss's Theorem asserts that there exists $v \in (\Lambda')^\bullet$ with $|q(v)| \leq |q|(v) < 2abc$, and since by construction $q(v) \equiv 0 \pmod{2abc}$, we must then have $q(v) = 0$, and thus we have found a solution with an explicit upper bound on the size of the coefficients.

On the other hand, if $n > 2$, then things work more strongly in our favor: there exists $v \in (\mathbb{Z}^3)^\bullet$ with $q(v) \equiv 0 \pmod{abcn}$ and $|q(v)| \leq 2^{\frac{1}{3}} n^{\frac{2}{3}} abc < nabc$, and we deduce that $q(v) = 0$ without having to compare q with the extremal form q_3. In fact, for sufficiently large n we can make the argument go through using Minkowski's Theorem 9.6 rather than the optimal constant γ_3. (I invite you to

calculate how large n needs to be for this.)

Cochrane and Mitchell take the former route: they show that one can take $n = 2$ and check the non-H-equivalence of $|q|_{\Lambda'}$ with the extremal form q_3. Let's check the second part first: put

$$M_{|q|} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix},$$

and let $A \in M_3(\mathbb{Z})$ be such that $\Lambda' = A\mathbb{Z}^3$. For this part of the computation we use only that $\det A = [\mathbb{Z}^3 : \Lambda'] = 2abc$. Then

$$\operatorname{disc} |q|_{\Lambda'} = \operatorname{disc}(|q|(Av)) = \det(A^t M_{|q|} A) = 4(abc)^3.$$

Now if $\lambda \in \mathbb{R}^{>0}$ is such that $|q|_{\Lambda'} \cong_{GL_3(\mathbb{Z})} \lambda q_3$, then

$$4(abc)^3 = \operatorname{disc} |q|_{\Lambda'} = \lambda^3 \operatorname{disc} \lambda q_3 = \frac{\lambda^3}{2},$$

and thus we must have

$$\lambda = 2abc.$$

From this it would follow that for all $(x, y, z) \in \Lambda'$ we would have

$$|q|(x, y, z) = |a|x^2 + |b|y^2 + |c|z^2 \equiv 0 \pmod{2abc},$$

whereas we also have

$$q(x, y, z) = |a|x^2 - |b|y^2 - |c|z^2 \equiv 0 \pmod{2abc},$$

so $2|a|x^2 \equiv 0 \pmod{2abc}$, and thus $x^2 \equiv 0 \pmod{bc}$, so $bc \mid x^2$. But it is easy to see that the x-coordinate of an element of Λ can be arbitrarily prescribed modulo bc, so this restriction on Λ' is only possible if $|bc| = 2$, so without loss of generality only if $b = -2$, $c = -1$. But then the equation is $ax^2 - 2y^2 - z^2 = 0$ and the congruence conditions imply that -2 is a square modulo a, so by our work on the binary form $x^2 + 2y^2$ we find that there is a solution with $x = 1$, and any such solution is *small* in the sense that $|x| \leq \sqrt{bc} = \sqrt{2}$, $|y| \leq \sqrt{|ac|} = \sqrt{a}$, $|z| \leq \sqrt{|ab|} = \sqrt{2a}$. So $|q|_\Lambda$ is not H-equivalent to q_3.

Finally, we must build the index 2 sublattice $\Lambda' \subset \Lambda$. The easiest case is when abc is odd: then we get Λ' by intersecting with the kernel of $q(x, y, z) \equiv x^2 + y^2 + z^2 = x + y + z : \mathbb{F}_2 \to \mathbb{F}_2$. When abc is even, we need to consider $q(x, y, z)$ modulo 4, and then (up to relaballing the variables) we have $q(x, y, z) \equiv 2x^2 + by^2 + cz^2 \pmod 4$ with b and c odd. By a modest amount of brute force one can find a codimension one – i.e., index 4 – subspace of $(\mathbb{Z}/4\mathbb{Z})^3$ on which q vanishes identically. For instance, when $a \equiv b \equiv 1 \pmod 4$, then $\langle (1, 1, 1),\ (1, 3, 3) \rangle$ is such a subspace. We leave the other cases to the reader.

Thus we have shown that if the necessary congruence conditions of Theorem 16.3 hold, there is $(x, y, z) \in (\mathbb{Z}^3)^\bullet$ such that

(i) $ax^2 + by^2 + cz^2 = 0$ and
(ii) $|a|x^2 + |b|y^2 + |c|z^2 < 2abc$.

Inequality (ii) immediately gives

$$|x| < \sqrt{2}\sqrt{|bc|},\ |y| < \sqrt{2}\sqrt{|ac|},\ |z| < \sqrt{2}\sqrt{|ab|}.$$

In fact, by combining (i) and (ii) we can deduce a sharper inequality. We state this curious fact in a slightly more general form due to L.M. Nunley [Nu10, Prop. 18].

Proposition 16.8. *Let $a_1, \ldots, a_N \in \mathbb{Z}^+$ be pairwise coprime and squarefree. Consider the following two norms on $\mathbb{R}^N$:*

$$|x| = |(x_1, \ldots, x_N)| = \max_i \frac{|x_i|}{\sqrt{a_1 \cdots a_{i-1} a_{i+1} \cdots a_N}},$$

$$||x|| = ||(x_1, \ldots, x_N)|| = \sqrt{a_1 x_1^2 + \ldots + a_N x_n^2}.$$

Suppose $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ satisfies

$$a_1 x_1^2 + \ldots + a_{N-1} x_{N-1}^2 - a_N x_N^2 = 0. \tag{78}$$

Then

$$||x|| = \sqrt{2a_1 \cdots a_N}|x|.$$

Proof. Since $x = (x_1, \ldots, x_N)$ satisfies (78) we have

$$a_1 x_1^2 + \ldots + a_{N_1} x_{N-1}^2 = a_N x_N^2,$$

and thus

$$\left(\sqrt{2a_1 \cdots a_N}|x|\right)^2 = 2a_1 \cdots a_N \left(\frac{x_N^2}{a_1 \cdots a_{N-1}}\right) = 2a_N x_N^2$$

$$= a_N x_N^2 + a_N x_N^2 = (a_1 x_1^2 + \ldots + a_{N-1} x_{N-1}^2) + a_N x_N^2 = ||x||^2.$$

□

Applying Proposition 16.8 to our nonzero $v = (x, y, z) \in \mathbb{Z}^3$ with $ax^2 + by^2 - |c|z^2 = 0$ and $ax^2 + by^2 + |c|z^2 \leq 2abc$, we find that since $||v|| \leq \sqrt{2abc}$, $|v| \leq 1$, and thus

$$|x| \leq \sqrt{|bc|},\ |y| \leq \sqrt{|ac|},\ |z| \leq \sqrt{|ab|}.$$

Summing up, we have completed the proof of the following result.

Theorem 16.9. *(Cochrane-Mitchell* [CoMi98]*) Let $a, b, c \in \mathbb{Z}$ with $a > 0$, $b, c < 0$ and abc squarefree. The following are equivalent:*
a) We have that $-ab$ is a square modulo c, $-ac$ is a square modulo b and $-bc$ is a square modulo a.
b) There exists a nonzero solution (x, y, z) to the Legendre equation

$$ax^2 + by^2 + cz^2 = 0.$$

*c) There exists a nonzero solution (x, y, z) to the Legendre equation which is **small** in the sense that $|q|(x, y, z) = |a|x^2 + |b|y^2 + |c|z^2 \leq 2abc$; or equivalently (among solutions to the Legendre equation) for which $|x| \leq \sqrt{bc}$, $|y| \leq \sqrt{ac}$, $|z| \leq \sqrt{ab}$.*

Question: Is there any $n > 2$ for which there exists an index $abcn$ sublattice Λ' of $\mathbb{Z}^3$ such that $q|_{\Lambda'} \equiv 0 \pmod{abcn}$?

16.5. **Nunley's Thesis.**

In [Nu10], Laura Nunley studied the **Extended Legendre Equation**

$$a_1x_1^2 + \ldots + a_{n-1}x_{n-1}^2 - a_nx_n^2,$$

with $a_1, \ldots, a_n \in \mathbb{Z}^+$, $a_1 \cdots a_n$ squarefree. She showed that the Cochrane-Mitchell argument goes through to give small solutions *provided* there exists a sublattice Λ of $\mathbb{Z}^n$ with $[\mathbb{Z}^n : \Lambda] = a_1 \cdots a_n$ and $q|_\Lambda \equiv 0 \pmod{a_1 \cdots a_n}$. Then she showed that – as long as $a_1 \cdots a_n > 2$; in the other cases she exhibits small solutions by easy arguments – such a sublattice Λ *does not exist*!

At the time, I found this negative result very surprising. In fact the idea that much of the meat of these GoN arguments is finding the magic sublattice of $\mathbb{Z}^n$ was one that became clear to us over the course of her thesis work, and it has informed much of the work in these notes and the corresponding VRG. Using some quadratic form theory, it is actually rather clear that these lattices cannot exist under the above hypotheses: namely they ensure that upon reducing modulo any odd prime p dividing $\operatorname{disc} q = a_1 \cdots a_n$, q becomes degenerate *but not degenerate enough*! Namely, modulo such a p q is the direct sum of a one-dimensional identically zero quadratic space $R(q)$ together with an $n-1$-dimensional nondegenerate quadratic space q'. A nondegenerate quadratic form q' over a field of characteristic not 2 cannot possibly have a totally isotropic subspace of dimension any larger than $\frac{\dim q'}{2}$, and equality occurs iff q' is a direct sum of hyperbolic planes. Thus, the dimension of the largest totally isotropic subspace of q is at most $1 + \frac{\dim q'}{2} = \frac{n+1}{2}$, so its codimension is at least $n - \frac{n+1}{2} = \frac{n-1}{2}$, which is greater than 1 for all $n > 3$.

Thus I am rather convinced that this GoN argument cannot be extended to prove the local-global principle for quadratic forms in more than three variables over $\mathbb{Q}$. **But** conceivably it can be extened in another way: the local-global principle holds not only over $\mathbb{Q}$ but – suitably formulated – for quadratic fields over an arbitrary **global field** K, i.e., a finite separable extension of either $\mathbb{Q}$ or $\mathbb{F}_p(t)$.

My feeling at the moment is that, after $\mathbb{Q}$, the most tractable case should be a rational function field $\mathbb{F}_q(t)$ (with q an odd prime power). In this regard we note that a linear forms theorem over function fields appears in §12.2.

17. GoN Applied to Diophantine Equations Over Number Fields

17.1. **Reminders on quadratic forms over number fields.**

For a number field K, we let $\mathbb{Z}_K$ be the ring of integers of K. Recall that a number field K with $[K : \mathbb{Q}] = d$ has r real embeddings and s conjugate pairs of complex embeddings, with $r + 2s = n$. We denote these embeddings by $|\cdot|_{\infty_i}$ for $1 \leq i \leq r+s$, it being understood that the first r are real and the last s are complex. An element $x \in K$ is **totally positive** if $\iota(x) > 0$ for every real embedding.

There is a canonical norm function $|\cdot| : \mathbb{Z}_K \to \mathbb{N}$ having the following basic properties: for all $x, y \in \mathbb{Z}_K$,

- $|x| = 0 \iff x = 0$.

- $|x| = 1 \iff x \in \mathbb{Z}_K^\times$.
- $|xy| = |x||y|$.

We extend the norm function to a function from K to $\mathbb{Q}^{\geq 0}$ by multiplicativity: $|\frac{a}{b}| = \frac{|a|}{|b|}$. (Clearly we still have $|x| = 0 \iff x = 0$. But for $x \in K$, $|x| = 1$ need not imply $x \in \mathbb{Z}_K^\times$.

Okay, but what is the definition of the norm function? It turns out that there are three useful definitions of one and the same norm function. Here we just give the deinitions. A proof of their equivalence is a good exercise for a student of algebraic number theory, or see

`http://math.uga.edu/~pete/GoNLinearForms.pdf`

where the equivalence is spelled out in gory detail.

First Definition: Of course put $|0| = 0$. For $x \in \mathbb{Z}_K^\bullet$,

$$|x| = \#R/(x).$$

Second Definition: For $x \in K$, put

$$|x| = \prod_{i=1}^{r+s} |x|_{\infty_i}.$$

Third Definition: For $x \in K$, put

$$|x| = |N_{K/\mathbb{Q}}(x)|.$$

Given a quadratic form $q(x) = q(x_1, \ldots, x_n)$ defined over K, for each real embedding ι of K we may extend scalars $\iota : K \hookrightarrow \mathbb{R}$ and get a real quadratic form. Thus it makes sense to describe a quadratic form as positive definite, positive semidefinite, indefinite, etc. with respect to any given real embedding ι. We say that q is **totally positive definite** if it is positive definite with respect to all real places.[21]

Exercise: Suppose $q(x) \cong a_1 x_1^2 + \ldots + a_n x_n^2$ with $a_1 \cdots a_n \neq 0$. Show that q is totally positive definite iff for $1 \leq i \leq n$, a_i is totally positive. In particular, for any $n \in \mathbb{Z}^+$, the sum of squares form $x_1^2 + \ldots + x_n^2$ is totally positive definite.

If $q(x) = q(x_1, \ldots, x_n)$ is totally positive definite, then it represents only totally positive elements of K. Conversely, it follows from the Hasse-Minowski theory that if $n \geq 4$, a totally positive definite form q K-represents all totally positive elements of K. In particular, if K has no real places then every quadratic form in $n \geq 4$ variables is universal.

Although the theory of quadratic forms over a number *field* K is every bit as complete and satisfactory as the special case $K = \mathbb{Q}$, the study of quadratic forms over the ring of integers $\mathbb{Z}_K$ of an arbitrary number field is much less developed than the corresponding theory over $\mathbb{Z}$.

[21] In particular, if K has no real places, then every quadratic form over K is trivially totally positive definite.

Let $q(t_1, \ldots, t_N) \in K[t_1, \ldots, t_n]$ be a quadratic form which is nondegenerate: $\operatorname{disc} q \neq 0$. Directly generalizing the $K = \mathbb{Q}$ case, we define the **Hermite invariant**

$$\gamma(q) = \frac{\inf_{x \in (\mathbb{Z}_K)^N \bullet} |q(x)|}{|\operatorname{disc} q|^{\frac{1}{N}}}.$$

We define the **Hermite constant**

$$\gamma_N(\mathbb{Z}_K) = \sup \gamma(q)$$

as q ranges over all nondegenerate N-ary quadratic forms. Similarly we define the **positive Hermite constant** $\gamma_N^+(\mathbb{Z}_K)$ by restricting the supremum to totally positive definite forms.

17.2. **Sums of Two Squares in Integral Domains.**

Let R be a domain with fraction field K. We say that R is **imaginary** if there exists $i \in K$ with $i^2 = -1$. Otherwise R is **non-imaginary**.

Lemma 17.1. *Suppose R is integrally closed in K. Then R is imaginary iff there exists $i \in R$ such that $i^2 = -1$.*

Exercise: Prove Lemma 17.1.

Lemma 17.2. *Suppose R has characteristic 2.*
a) Then R is imaginary.
b) For any $n \in \mathbb{Z}^+$, an element of R is a sum of n squares iff it is a square.

Exercise: Prove Lemma 17.2.

Lemma 17.3. *Let $\mathbb{F}$ be a finite field of order p^a (with p a prime number). Then $\mathbb{F}$ is imaginary iff $p = 2$, $p \equiv 1 \pmod 4$ or a is even.*

Exercise: Prove Lemma 17.3.

If R is non-imaginary, put $R[i] := R[t]/(t^2 + 1)$. Then every $z \in R[i]$ has a unique expression of the form $a + bi$ with $a, b \in R$. The fraction field of $R[i]$ is $K[i] = K[t]/(t^2+1)$, $K[i]/K$ is a separable quadratic extension of fields and $R[i]/R$ is an integral extension of domains. We denote by $z \mapsto \overline{z}$ the unique nontrivial automorphism of $K[i]$: explicitly $\overline{x + iy} = x - iy$. This automorphism stabilizes $R[i]$. Let $N : K[i] \to K$ denote the norm map:

$$N(x + iy) = (x + iy)(\overline{x + iy}) = (x + iy)(x - iy) = x^2 + y^2.$$

This is a multiplicative map which restricts to a multiplicative map $R[i] \to R$.

Let x be an element of a domain R. We say that x is **a sum of two squares up to a unit** if there exist $a, b \in R$ and $u \in R^\times$ such that $x = u(a^2 + b^2)$. For a prime element p of R, let F_p be the fraction field of the domain R/pR. We say p is imaginary (resp. nonimaginary) if F_p is imaginary (resp. nonimaginary).

The following is a small result I proved in the summer of 2010. I later learned that similar results (including a proof of part c)) were attained quite a while ago by Choi, Lam, Resnick and Rosenberg [CLRR80].

Theorem 17.4. *Let R be a non-imaginary domain, and let $\rho \in R$ be a prime element. Consider the following two assertions:*
(i) ρ is a sum of two squares up to a unit.
(ii) The field F_ρ is imaginary.
a) In all cases (i) $\implies$ (ii).
b) Suppose $R[i]$ is a UFD. Then (ii) $\implies$ (i).
c) Suppose R is a UFD. Then an element $f \in R \setminus \{0\}$ is a sum of two squares up to a unit iff for every prime element ρ such that F_ρ is non-imaginary, $\mathrm{ord}_\rho(f)$ *is even.*

Proof. a) Let ρ be a prime element of R; suppose there are $a, b \in R$, $u \in R^\times$ with

$$\rho = u(a^2 + b^2). \tag{79}$$

We will write the composite homomorphism $R \to R/\rho R \to F_\rho$ as $z \in R \mapsto \underline{z} \in F_\rho$. Applying this map to both sides of (79), we get

$$0 = \underline{u}(\underline{a}^2 + \underline{b}^2).$$

Since ring homomorphisms send units to units, $\underline{u} \in F_\rho^\times$, and thus

$$\underline{a}^2 + \underline{b}^2 = 0.$$

If $\underline{a} = 0$, then $\rho \mid a$ and thus from (79), $\rho \mid b$. Writing $a = \rho A$, $b = \rho B$ for $A, B \in R$, substituting into (79) and cancelling ρ, we get $1 = u\rho(A^2 + B^2)$ and thus ρ is a unit, contradiction. So $\underline{a} \in F_\rho^\times$ and F_ρ is imaginary.
b) Suppose that F_ρ is imaginary: there exists $\underline{r} \in F_\rho$ with $\underline{r}^2 + 1 = 0$. Then there are $x, y \in R \setminus \rho R$ such that $\underline{r} = \frac{\underline{x}}{\underline{y}}$, so $\left(\frac{\underline{x}}{\underline{y}}\right)^2 + 1 = \rho z$ for $z \in R$. Clearing denominators and working in $R[i]$ we get

$$y^2 \rho z = x^2 + y^2 = (x + iy)(x - iy).$$

We claim that ρ is *not* a prime element of $R[i]$. Indeed, if it were, then since $\rho \mid (x + iy)(x - iy)$, we would have $\rho \mid (x \pm iy)$, i.e., $\frac{x}{\rho} \pm i\frac{y}{\rho} \in R[i]$. But this implies $\rho \mid y$, contradiction. Since ρ is a nonprime element in the UFD $R[i]$, there exist nonzero nonunit elements $\alpha, \beta \in R[i]$ such that $\rho = \alpha\beta$. Taking norms gives

$$\rho^2 = N(\rho) = N(\alpha\beta) = N(\alpha)N(\beta).$$

Now the prime element ρ of R divides $N(\alpha)N(\beta)$, so it divides one of the factors. It is no loss of generality to assume that $\rho \mid N(\alpha)$:

$$\rho = \left(\frac{N(\alpha)}{\rho}\right) N(\beta).$$

Thus either $\rho \mid \frac{N(\alpha)}{\rho}$ or $\rho \mid N(\beta)$. If the former occurs, then we get an equation

$$1 = \left(\frac{N(\alpha)}{\rho^2}\right) N(\beta),$$

so $N(\beta) \in R^\times$. But this cannot be: β is not a unit in $R[i]$ and $N(\beta) = \beta\overline{\beta}$, so $N(\beta)$ is not a unit in $R[i]$ and thus, *a fortiori*, not a unit in R. Thus we have

$$1 = \left(\frac{N(\alpha)}{\rho}\right)\left(\frac{N(\beta)}{\rho}\right),$$

so there are $u_1, u_2 \in R^\times$ such that $\rho = u_1 N(\alpha) = u_2 N(\beta)$. Put $\alpha = a + bi$; then $N(\alpha) = a^2 + b^2$ and $\rho = u_1(a^2 + b^2)$.

c) Suppose R is a UFD, and let $f \in R \setminus \{0\}$; we may assume $f \notin R^\times$. By unique factorization, we may write $f = f_1^2 f_2$ with f_2 not divisible by the square of any prime element of R. Write $f_2 = p_1 \cdots p_k$ a product of nonassociate prime elements.

Suppose that $\text{ord}_p(f)$ is even for all imaginary prime elements p; then each prime element p_i dividing f_2 is imaginary. By part a), for all i there exist $a_i, b_i \in R$ an $u_i \in R^\times$ such that $p_i = u_i(a_i^2 + b_i^2)$. Moreover, the multiplicativity of the norm map $N : \mathbb{Z}[a,b,c,d][i] \to \mathbb{Z}[a,b,c,d]$ gives rise to the identity

$$\left(a^2 + b^2\right)\left(c^2 + d^2\right) = (ac - bd)^2 + (ad + bc)^2 \tag{80}$$

(alternate proof: direct calculation!), which shows that in any integral domain, the set of sums of two squares is closed under multiplication, so all in all we have

$$f = u_1 \cdots u_n(A^2 + B^2) = u(A^2 + B^2).$$

Thus f is a sum of two squares up to a unit.

Conversely, suppose there exist $a, b \in R$ and $u \in R^\times$ such that $f = u(a^2 + b^2)$, and let p be a non-imaginary prime divisor of f. Reducing modulo p gives

$$0 = \underline{u}(\underline{a}^2 + \underline{b}^2),$$

and thus, as above,

$$(\underline{a})^2 + (\underline{b})^2 = 0.$$

But in the non-imaginary field F_p this forces $\underline{a} = \underline{b} = 0$, i.e., there are $A, B \in R$ such that $a = pA$, $b = pB$ and thus $f = p^2 u(A^2 + B^2)$. So $\frac{f}{p^2}$ is again a sum of two squares up to a unit. We may continue this process, but not infinitely many times, since $\text{ord}_p(f)$ is finite. Eventually it must end, showing that $\text{ord}_p(f)$ is even. □

Theorem 17.4 represents a sort of "rival" to Geometry of Numbers methods. Here are some of its immediate applications.

Corollary 17.5. *(Full Two Squares Theorem) A nonzero integer n is a sum of two integer squares iff $n \geq 0$ and $\text{ord}_p(n)$ is even for all $p \equiv 3 \pmod 4$.*

Proof. Since $\mathbb{Z}$ is a nonimaginary domain such that $\mathbb{Z}[i]$ is a UFD (indeed the standard norm function is Euclidean), Theorem 17.4b) applies. By Lemma 17.3, a prime p in $\mathbb{Z}$ is imaginary iff $p \equiv 3 \pmod 4$, so Theorem 17.4b) says that $n \in \mathbb{Z} \setminus \{0\}$ is a sum of two squares up to a unit iff $\text{ord}_p(n)$ is even for all $p \equiv 3 \pmod 4$. The units in $\mathbb{Z}$ are ± 1, so if $n = u(x^2 + y^2)$ we have $u = 1 \iff n > 0$. □

Corollary 17.6. *(Two Squares Theorem in $\mathbb{R}[t]$) For a polynomial $f \in \mathbb{R}[t]$, TFAE:*
(i) There exist $a, b \in \mathbb{R}[t]$ such that $f(t)^2 = a(t)^2 + b(t)^2$.
(ii) For all $t \in \mathbb{R}$, $f(t) \geq 0$.

Exercise: Prove Corollary 17.6 using Theorem 17.4. (Hint: recall that every polynomial with real coefficients factors into a product of linear factors and irreducible quadratic factors. Show that the first type of prime element is nonimaginary and the second type of prime element is imaginary. Use high school algebra to relate the factorization of $p(t)$ into prime elements to the condition $f(t) \geq 0$ for all $t \in \mathbb{R}$.)

Corollary 17.7. *(Leahey's Theorem)*
Let $\mathbb{F}$ be a finite field of order p^a, and let $R = \mathbb{F}[t]$.
a) If $p = 2$, then $f \in R$ is a sum of two squares iff it is a square.
b) If $p^a \equiv 1 \pmod 4$, then every $f \in R$ is a sum of two squares.

c) If $p^a \equiv 3 \pmod 4$, then $f \in R^\bullet$ is a sum of two squares iff $\mathrm{ord}_p(f)$ is even for every irreducible polynomial p of odd degree.

Exercise: Use Theorem 17.4 to prove Corollary 17.7.

17.3. Sums of two squares in $\mathbb{Z}_K$, $K = \mathbb{Q}(\sqrt{5})$.

The goal of this section is to prove the following result of J.I. Deutsch [De02].

Let $\epsilon = \frac{1+\sqrt{5}}{2}$, the fundamental unit of $\mathbb{Z}_K$. Also $\mathbb{Z}_K = \mathbb{Z}[\epsilon] = \mathbb{Z} \cdot 1 \oplus \mathbb{Z} \cdot \epsilon$.

Theorem 17.8. *(Deutsch [De02]) Let $K = \mathbb{Q}(\sqrt{5})$. For a prime element ρ of the PID $\mathbb{Z}_K$, the following are equivalent:*
(i) -1 is a square in $\mathbb{Z}_K/\rho$.
(ii) ρ is a sum of two squares in $\mathbb{Z}_K$, up to a unit: there exist $x, y \in \mathbb{Z}_K$ and $u \in \mathbb{Z}_K^\times$ such that $x^2 + y^2 = u\rho$.

Proof. Step 0: Since $\sqrt{-1} \notin K$, $\mathbb{Z}_K$ is not imaginary and Theorem 17.4a) gives (ii) $\implies$ (i).[22] For the proof proper, let p be the prime of $\mathbb{Z}$ lying below ρ, i.e., the characteristic of $\mathbb{Z}_K/\rho$. Reasonably enough, the argument will treat separately the three possibilities for the splitting of p in K.
Step 1: Suppose p ramifies in $\mathbb{Z}_K$. Since the discriminant of $\mathbb{Z}_K$ is 5, this holds iff $p = 5$, in which case $\rho = v\sqrt{5}$ for $v \in \mathbb{Z}_K^\times$. In this case we may proceed by brute force: $\epsilon\sqrt{5} = 1^2 + \epsilon^2$, so $\rho = (\frac{v}{\epsilon})(\epsilon\sqrt{5})$ is a sum of two squares up to a unit.
Step 2a: Suppose p splits in $\mathbb{Z}_K$. Let $\lambda \in \mathbb{Z}_K$ be such that $\lambda^2 \equiv -1 \pmod \rho$. We consider the lattice $\Lambda_\rho = M_\rho \mathbb{Z}^4$, with

$$M_\rho = \begin{pmatrix} 1 & \epsilon & 0 & 0 \\ 1 & \overline{\epsilon} & 0 & 0 \\ \lambda & \lambda\epsilon & \rho & \epsilon\rho \\ \overline{\lambda} & \overline{\lambda\epsilon} & \overline{\rho} & \overline{\epsilon\rho}. \end{pmatrix}$$

We have $\mathrm{Covol}\,\Lambda_\rho = |\det M_\rho| = |5\rho\overline{\rho}| = 5p$. By Minkowski's Convex Body Theorem, the R-ball in $\mathbb{R}^4$ contains a nonzero point of Λ_ρ iff

$$\mathrm{Vol}(B_4(r)) = \frac{\pi^2 R^4}{2} \geq 2^4 \cdot (5p)$$

$$\iff R^2 \geq \frac{\sqrt{160p}}{\pi} \leq 4.1\sqrt{p}.$$

Now general element of Λ_ρ is of the form

$$(\alpha, \overline{\alpha}, \lambda\alpha + \mu\rho, \overline{\lambda\alpha + \mu\rho})$$

for arbitrary $\alpha, \mu \in \mathbb{Z}_K$. In particular any element of Λ_ρ is of the form $(\alpha, \overline{\alpha}, \beta, \overline{\beta})$ for some $\alpha, \beta \in \mathbb{Z}_K$ such that $\alpha^2 + \beta^2 \equiv 0 \pmod \rho$ (this follows from the relation $\lambda^2 \equiv -1 \pmod \rho$ exactly as for the Two Squares Theorem over $\mathbb{Z}$). Therefore, taking $0 \neq v \in \Lambda_\rho$ with $v \cdot v \leq 4.1\sqrt{p}$, we get $v = (\alpha, \beta, \overline{\alpha}, \overline{\beta})$ with

$$\alpha^2 + \beta^2 = \kappa\rho,$$

and since $\alpha^2 + \overline{\alpha}^2 + \beta^2 + \overline{\beta}^2 < 4.1\sqrt{p}$,

$$\kappa\rho + \overline{\kappa}\rho < 4.1\sqrt{p}.$$

[22] In [De02], Deutsch actually proves the implication (i) $\implies$ (ii) only. This is certainly the more interesting direction, and probably he was aware of (ii) $\implies$ (i) as well.

Since $\kappa\rho$ is a sum of squares, it is totally positive. The AGM Inequality gives

$$2\sqrt{(\kappa\rho)(\overline{\kappa\rho})} \leq \kappa\rho + \overline{\kappa\rho} < 4.1\sqrt{p},$$

so

$$(\kappa\rho)(\overline{\kappa\rho}) \leq \frac{4.1^2}{4}p$$

and thus

$$|\kappa\overline{\kappa}| \leq \lfloor \frac{4.1^2}{4} \rfloor = 4.$$

Step 2b: Write $\kappa = a + b\epsilon$ for $a, b \in \mathbb{Z}$. We have

$$\kappa\overline{\kappa} = (a+b\epsilon)(a+b\overline{\epsilon}) = a^2 + ab(\epsilon+\overline{\epsilon}) + b^2\epsilon\overline{\epsilon} = a^2 + ab - b^2$$

$$\equiv -4a^2 - 4ab - b^2 \equiv -(2a+b)^2 \pmod 5.$$

Since the squares mod 5 are $0, \pm 1$, it follows that $N(\kappa) = \kappa\overline{\kappa}$ cannot equal $\pm 2, \pm 3$. Since $v \neq 0$, $N(\kappa) \neq 0$, so the alternatives are $N(\kappa) = \pm 1$ – so κ is a unit and we are done – or $N(\kappa) = \pm 4$. Since 2 is inert in $\mathbb{Z}_K$, if $\kappa\overline{\kappa} = \pm 4$, then $2 \mid \kappa$ or $2 \mid \overline{\kappa}$, and if one of these divisibilities holds they both hold. So we may wite $\kappa = 2\mu$ with $N(\mu) = \pm 1$ so $\mu \in \mathbb{Z}_K^\times$ and thus $\alpha^2 + \beta^2 = \kappa\rho = 2\mu\rho$. Therefore

$$0 \equiv \alpha^2 + \beta^2 \equiv (\alpha+\beta)^2 \pmod 2.$$

Since 2 is a prime element, this gives $2 \mid \alpha+\beta$. Of course $\mod 2$ we have $\alpha^2+\beta^2 = \alpha^2-\beta^2$, so also $2 \mid (\alpha-\beta)$. We may therefore divide through to get a representation of $\mu\rho$ as a sum of two squares:

$$\left(\frac{\alpha+\beta}{2}\right)^2 + \left(\frac{\alpha-\beta}{2}\right)^2 = \frac{\alpha^2+\beta^2}{2} = \mu\rho.$$

Step 3: Finally, we suppose that p is inert in $\mathbb{Z}_K$, i.e., is a prime element of $\mathbb{Z}_K$: in terms of quadratic symbolswe are assuming $(\frac{5}{p}) = -1$. If $p = 2$ then we are okay: $2 = 1^2 + 1^2$. Indeed if $p \equiv 1 \pmod 4$ we are also okay because then – as we have already proven by GoN methods – p is already a sum of two integral squares. We need then to consider the case $p \equiv 3 \pmod 4$. Note that in this case we have

$$\left(\frac{-5}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{5}{p}\right) = (-1)\cdot(-1) = 1.$$

So Theorem 15.13 applies: there are $x, y \in \mathbb{Z}$ and $1 \leq k \leq 2$ such that

$$kp = x^2 + 5y^2 = x^2 + (\sqrt{5}y)^2.$$

If $k = 1$, then the above identity shows that we have represented p as a sum of 2 squares in $\mathbb{Z}_K$. (In fact the equation $p = x^2+5y^2$ is impossible if $p \equiv 3 \pmod 4$, as one sees by reducing modulo 4. We proceed on to the remaining case.) Reducing $2p = x^2 + 5y^2$ modulo 2 shows that x and y have the same parity, and we may thus write $x = y + 2a$, $a \in \mathbb{Z}$. Then

$$(a+y\epsilon)^2 + (a+y\overline{\epsilon})^2 = 2a^2 + 2ay(\epsilon+\overline{\epsilon}) + y^2(\epsilon^2+\overline{\epsilon}^2) = 2a^2 + 2ay + 3y^2$$

$$= \frac{4a^2+4ay+6y^2}{2} = \frac{(2a+y)^2+5y^2}{2} = \frac{x^2+5y^2}{2} = \frac{2p}{2} = p.$$

□

Remark: The implication (i) $\implies$ (ii) would follow from Theorem 17.4b) if the ring $S = \mathbb{Z}_K[\sqrt{-1}]$ were a UFD. In fact S is not a UFD, but for a rather superficial reason: it is not integrally closed in $L = \mathbb{Q}(\sqrt{-1}, \sqrt{-5})$. The following MAGMA code shows that S is a proper subring of $\mathbb{Z}_L$ and that $\mathbb{Z}_L$ *is* a UFD.

```
> K1<i> := QuadraticField(-1);
> i^2;
-1
> K2<s> := QuadraticField(-5);
> s^2;
-5
> L := Compositum(K1,K2);
> P := MinimalPolynomial(i+s);
> P;
.1^4 + 12 * .1^2 + 16
> Discriminant(P);
1638400
> Factorization(1638400);
[ <2, 16>, <5, 2> ]
ClassNumber(L);
1
> Factorization(Discriminant(MaximalOrder(L)));
[ <2, 4>, <5, 2> ]
```

So $\mathbb{Z}_L$ is a UFD and $S = \mathbb{Z}_K[\sqrt{-1}]$ sits inside $\mathbb{Z}_L$ with index 2^6. S is close to satisfying the hypotheses of Theorem 17.5b), and I suspect that by comparing S with the UFD $\mathbb{Z}_L$ one should be able to prove (i) $\implies$ (ii) of Deutsch's Theorem.

17.4. **Sums of Squares in $\mathbb{Z}[i]$.**

Let $\mathbb{Z}[i] = \mathbb{Z}[\sqrt{-1}]$ be the ring of Gaussian integers. Which elements of $\mathbb{Z}[i]$ are sums of two squares? There is an interesting arithmetic limitation.

Lemma 17.9. *Let $z = x + yi \in \mathbb{Z}[i]$.*
a) If there are $x_1, \ldots, x_n \in \mathbb{Z}[i]$ with $z = x_1^2 + \ldots + x_n^2$, then y is even.
b) If $z = u + vi$ with $u \equiv v \pmod 2$, then $1 \pm i \mid z$.

Proof. Left as an easy exercise. □

Thus we should ask which Gaussian integers of the form $a+2bi$ are sums of squares.

Theorem 17.10. *For $a, b \in \mathbb{Z}$, let $z = a + 2bi \in \mathbb{Z}[i]$.*
a) (Niven [Ni40]*) The following are equivalent:*
(i) Either $a \not\equiv 2 \pmod 4$ or $b \equiv 0 \pmod 2$.
(ii) There are $u, v \in \mathbb{Z}[i]$ such that $z = u^2 + v^2$.
b) (Niven-Joly [Ni40] [Jo70]*) There are $a, b, c \in \mathbb{Z}[i]$ such that $z = a^2 + b^2 + c^2$.*

Proof. a) ([Le65]) (i) $\implies$ (ii): We'll explicitly write z as a sum of two squares. Case 1: Suppose a is odd. Then

$$z = a + 2bi = \left(\frac{z+1}{2}\right)^2 + i^2 \left(\frac{z-1}{2}\right)^2.$$

Case 2: Suppose $a = 2a'$ with a' odd and b is even. Put $z' = a' + bi$. Then

$$z = \left(\frac{z'+i}{1+i}\right)^2 + i^2\left(\frac{z'-i}{1+i}\right)^2.$$

Case 3: Suppose $a = 4a'$ and b is odd. Put $z' = b - 2a'i$. Then

$$z = \left(\frac{z'+1}{1-i}\right)^2 + i^2\left(\frac{z'-1}{1-i}\right)^2.$$

Case 4: Suppose $a = 4a'$ and $b = 2b'$. Put $z' = 2a' + 2b'i$. Then

$$z = \left(\frac{z'+2}{2}\right)^2 + i^2\left(\frac{z'-2}{2}\right)^2.$$

(ii) $\implies$ (i): Suppose

$$a + 2bi = (x + yi)^2 + (z + wi)^2$$

and $a = 2a'$ with $a' \equiv 1 \pmod 2$ and $b \equiv 1 \pmod 2$. Then

$$2(a' + bi) = a + 2bi = (x + yi)^2 + (z + wi)^2 = x^2 - y^2 + z^2 - w^2 + 2(xy + zw)i,$$

so

$$a = 2a' = x^2 - y^2 + z^2 - w^2, \tag{81}$$

$$b = xy + zw. \tag{82}$$

Since b is odd, (82) implies that exactly two or exactly three of x, y, z, w are odd, and (81) rules out the latter possibility. Furthermore, (82) implies that either ($x \equiv y \equiv 1 \pmod 2$ and $z \equiv w \equiv 0 \pmod 2$) or ($x \equiv y \equiv 0 \pmod 2$ and $z \equiv w \equiv 1 \pmod 2$). Either way, $4 \mid x^2 - y^2 + z^2 - w^2$, contradicting the hypothesis that a' is odd.
b) Let $z = a + 2bi \in \mathbb{Z}[i]$. If $b \equiv 0 \pmod 2$, then by part a) z is a sum of two squares. If $b \equiv 1 \pmod 2$, then by part a) there are $x, y \in \mathbb{Z}[i]$ such that

$$z - (1 + i)^2 = a + 2(b - 1)i = x^2 + y^2,$$

so

$$z = x^2 + y^2 + (1 + i)^2.$$

□

Exercise: Show: $z = a+2bi$ is a sum of two squares in $\mathbb{Z}[i]$ iff $(1+i)^3 \nmid z$ or $(1+i)^4 \mid z$.

This exercise is used to give another simple proof of Theorem 17.10 in [Wi73]. G. Pall gave a formula for the number of representations of $z \in \mathbb{Z}[i]$ by $x^2 + y^2$ [Pa51]. Williams gave a simple, short proof of Pall's formula in [Wi71].

Exercise: a) Show that the form $x^2 + y^2 + z^2 + iw^2$ is universal over $\mathbb{Z}[i]$: i.e., represents every Gaussian integer.
b) Show that there are infinitely many nonisomorphic diagonal quaternary universal forms over $\mathbb{Z}[i]$.
c)* Is there an anisotropic diagonal quaternary universal form over $\mathbb{Z}[i]$?

Exercise (F. Sidokhine): Show that $x^2 + y^2 + iz^2 + iw^2$ is universal over $\mathbb{Z}[i]$.

17.5. **Hermite constants in number fields.**

Theorem 17.11. *(Icaza* [Ic97]*) Let $K/\mathbb{Q}$ be a totally real of degree d. For all $n \in \mathbb{Z}^+$,*

$$\gamma_N^+(\mathbb{Z}_K) \leq 4^d V_N^{\frac{-2d}{N}} |d(K)|,$$

where $V_N = \frac{2\pi^{\frac{N}{2}}}{N\Gamma(\frac{N}{2})}$ is the volume of the unit ball in Euclidean N-space.

Proof. Let $\sigma_1, \ldots, \sigma_d : K \hookrightarrow \mathbb{R}$ be the real embeddings. Let $q \in Q(K,N)$ be totally positive. For $1 \leq i \leq d$, we put $q_i = \sigma_i(q_i) \in \mathbb{R}[t_1, \ldots, t_n]$: we apply σ_i to each coefficient of q_i, getting a positive real form. Then for all $x \in K^N$,

$$|q(x)| = \prod_{i=1}^d |q_i(x)| = \prod_{i=1}^d q_i(x)$$

and similarly

$$|\operatorname{disc} q| = |N_{K/\mathbb{Q}} \operatorname{disc} q| = \prod_{i=1}^d \operatorname{disc} q_i,$$

so

$$\gamma(q) = \inf_{x \in \mathbb{Z}_K^N \bullet} \prod_{i=1}^d \frac{q_i(x)}{(\operatorname{disc} q_i)^{\frac{1}{N}}}.$$

Each q_i, being a positive real form, defines a Minkowski functional with level sets

$$\Omega(q_i, R) = \{x \in \mathbb{R}^N \mid q_i(x) \leq R^2\}$$

which are ellipsoids. For $y_1, \ldots, y_d \in \mathbb{R}^N$ we also define

$$Q : \mathbb{R}^{dN} \to \mathbb{R},\ Q(y_1, \ldots, y_d) = \max_{1 \leq i \leq d} q_i(y_i).$$

The associated level sets are **polyellipsoids**, i.e., Cartesian products of ellipsoids:

$$\Omega(Q, R) = \{x \in \mathbb{R}^N \mid Q \leq R^2\} = \prod_{i=1}^d \Omega(q_i, R).$$

By Lemma 9.5 we have

$$\operatorname{Vol} \Omega(Q,R) = \prod_{i=1}^d \operatorname{Vol} \Omega(q_i, R) = \prod_{i=1}^d \frac{V_N}{\sqrt{\operatorname{disc} q_i}} R^N = \frac{V_N^d R^{dN}}{\sqrt{|\operatorname{disc} q|}}.$$

Let $\Lambda = \mathbb{Z}_K^N \subset \mathbb{R}^{dN}$; Λ is the Cartesian product of N copies of the usual lattice $\mathbb{Z}_K$ in $\mathbb{R}^d$, so by Proposition 11.1, $\operatorname{Covol} \Lambda = \sqrt{|d(K)|}^N$. We choose R such that

$$\frac{V_N^d R^{dN}}{\sqrt{|\operatorname{disc} q|}} = \operatorname{Vol} \Omega(Q,R) = 2^{dN} \operatorname{Covol} \Lambda = 2^{dN} |d(K)|^{\frac{N}{2}}.$$

Thus

$$R = 2|d(K)|^{\frac{1}{2d}} |\operatorname{disc} q|^{\frac{1}{2dN}} V_N^{\frac{-1}{N}}.$$

By MCBT there is $v \in \Omega(Q,R) \cap \Lambda^\bullet$. Then $q_i(v) \leq R^2$ for all i, so

$$|q(v)| = \prod_{i=1}^d q_i(v) \leq R^{2d} = 4^d |d(K)| |\operatorname{disc} q|^{\frac{1}{N}} V_N^{\frac{-2d}{N}}$$

and thus

$$\gamma(q) \leq \frac{|q(v)|}{|\operatorname{disc} q|^{\frac{1}{N}}} \leq 4^d V_N^{\frac{-2d}{N}} |d(K)|.$$

□

18. Geometry of Numbers Over Function Fields

18.1. **No, seriously.**

It sounds weird, but give it a chance.

If you don't know what a "function field" is, here is a good example of one: let k be any field, and let $K = k(t)$ be the field of rational functions with coefficients in k. Inside K we have $R = k[t]$, the polynomial ring.

Algebraists and number theorists have long appreciated the following analogy:

$$\mathbb{Z} : \mathbb{Q} :: k[t] : k(t).$$

In particular $\mathbb{Z}$ and $k[t]$ are both principal ideal domains. In fact, they are both principal ideal domains because they admit a Euclidean norm function $|\ |$: for $\mathbb{Z}$ it is the usual Archimedean absolute value; for $k[t]$ we can fix any real number $e > 1$ and put $|f| = e^{\deg f}$.

The analogies become sharper when $k = \mathbb{F}_q$ is a finite field...

18.2. **Tornheim's Linear Forms Theorem.**

In 1941, L. Tornheim gave an analogue of Minkowski's Linear Forms theorem in the function field case: let k be any field, let $R = k[t]$ and let $K = k(t)$.

First let's try to set this up: in the usual linear forms theorem we have a matrix $C = (c_{ij}) \in M_N(\mathbb{R})$ and a lattice $\Lambda \subset \mathbb{R}^N$. But as we saw in the proof, taking an arbitrary lattice does not actually result in a more general theorem, so let's restrict to $\Lambda = \mathbb{Z}^N$. Then the theorem says that for any $\epsilon_1, \ldots, \epsilon_N \in \mathbb{R}^{>0}$ such that

$$|\det C| \leq \prod_{i=1}^N \epsilon_i,$$

there exists $x \in (\mathbb{Z}^N)^\bullet$ such that for all $1 \leq i \leq N$,

$$|L_i(x)| = |\sum_{j=1}^N c_{ij} x_j| \leq \epsilon_i. \tag{83}$$

Just for kicks, let's consider the equivalent version of the Linear Forms Theorem resulting from taking the logarithm of both sides of (83):

$$\log |L_i(x)| \leq \sum_{i=1}^N \log \epsilon_i.$$

Here we take the convention that $\log 0 = -\infty$. And let's touch this up a little bit: since the ϵ_i's are arbitrary positive numbers, the $\log \epsilon_i$'s are arbitrary real numbers, so we can restate the theorem as follows.

Theorem 18.1. *(Massaged Minkowski's Linear Forms Theorem) Let $C = (c_{ij}) \in M_N(\mathbb{R})$; for $1 \leq i \leq N$, put $L_i(x) = \sum_{j=1}^N c_{ij}x_j$. Let $e_1, \ldots, e_N \in \mathbb{R}$ be such that*

$$\log |\det C| \leq \sum_{i=1}^N e_i.$$

Then there is $x \in (\mathbb{Z}^N)^\bullet$ such that for all $1 \leq i \leq N$,

$$\log |L_i(x)| \leq e_i.$$

Now it is straightforward to write down a function field analogue: we replace $\mathbb{Z}$ by $R = k[t]$, we replace $\mathbb{R}$ by $K = k(t)$, and – most crucially of all – we replace the function $\log | \ | : \mathbb{R} \to [-\infty, \infty)$ by $\deg : k(t) \to [-\infty, \infty)$.

Lemma 18.2. *Let k be a field and put $R = k[t]$. Let $C \in M_n(R) \cap \mathrm{GL}_n(K)$, and let $\Lambda = CR^n$. Then Λ is an R-submodule of R^n, so we may form the quotient R-module R^n/Λ. Then*

$$\dim_k(R^n/\Lambda) = \deg \det C.$$

Proof. The R-module R^n/Λ is unchanged if we replace C by ACB for any $A, B \in \mathrm{GL}_n(R)$. Since R is a PID, **Smith Normal Form** applies: there are $A, B \in \mathrm{GL}_n(R)$ such that ACB is a diagonal matrix with diagonal entries $d_1, \ldots, d_n \in R^\bullet$. Since $A, B \in \mathrm{GL}_n(R)$, $(\det C)R = (\det ACB)R$, and thus

$$\deg \det C = \deg \det ABC = \deg \prod_{i=1}^n d_i.$$

Moreover, $\Lambda = \bigoplus_{i=1}^n d_i R$, $R^n/\Lambda = \bigoplus_{i=1}^n R/(d_i)$, so

$$\dim_k R^n/\Lambda = \dim_k \bigoplus_{i=1}^n R/(d_i) = \sum_{i=1}^n \deg d_i = \deg \prod_{i=1}^n d_i = \deg \det C.$$

□

Theorem 18.3. *(Tornheim* [To41]*) Let k be a field; let $C = (c_{ij}) \in M_n(k[t])$ with $\det C \neq 0$. For $1 \leq i \leq n$, put $L_i(x) = \sum_{j=1}^n c_{ij}x_j$. Let $e_1, \ldots, e_n \in \mathbb{N}$ be such that*

$$\deg \det C < \sum_{i=1}^n (e_i + 1). \tag{84}$$

Then there exists $x \in (R^N)^\bullet$ such that for all $1 \leq i \leq N$,

$$\deg L_i(x) \leq e_i.$$

Proof. The key idea is to work backwards, i.e., with the inverse transformation. Let $R = k[t]$, $K = k(t)$, consider the linear transformation

$$L : K^n \to K^n, x \mapsto Cx,$$

and put $\Lambda = L(R^n) \subset K^N$. Since $\det C \neq 0$, we have

$$L^{-1} : K^n \to K^n.$$

By definition of Λ,

$$L^{-1}|_\Lambda : \Lambda \xrightarrow{\sim} R^n.$$

Let

$$\mathcal{B} = \{(x_1, \ldots, x_n) \in R^n \mid \forall 1 \leq i \leq n, \ \deg x_i \leq e_i\}.$$

Thus $\mathcal{B}$ is a k-subspace of R^n with

$$\dim_k \mathcal{B} = \sum_{i=1}^{n} (e_i + 1).$$

By Lemma 18.2, $\dim_k R^n/\Lambda = \deg \det C$. Then (84) can be restated as

$$\dim_k \mathcal{B} > \operatorname{codim}_k \Lambda.$$

Thus there is a nonzero vector $y \in \Lambda \cap \mathcal{B}$. Taking $x = L^{-1}y$ does the job. □

Exercise: a) Show that Theorem 18.3 holds verbatim for matrices $C \in \mathrm{GL}_n(k(t))$. b) State and prove a version of Theorem 18.3 for $C \in \mathrm{GL}_N(k((t)))$.

There is a version of Lemma 18.2 valid over any Dedekind domain R. In this case, R^n/Λ is a finite length R-module, and so admits a **Serre invariant** $\chi(R^n/\Lambda)$. Indeed, let M be a finite length R-module and consider a composition series

$$0 = M_0 \subset M_1 \subset \ldots \subset M_N = M.$$

Each successive quotient M_{i+1}/M_i is a simple R-module hence of the form $R/\mathfrak{p}_i$ for some maximal ideal $\mathfrak{p}_i$ of R. We put $\chi(M) = \prod_{i=1}^{N} \mathfrak{p}_i$. The Jordan-Hölder Theorem implies that $\chi(M)$ is independent of the chosen composition series.

Exercise: With notation as above, show that $\chi(R^n/\Lambda) = (\det C)R$. (Suggestion: reduce to the case in which R is a DVR and use Smith Normal Form.)

Suppose now we have a function $|\cdot|$ which associates to each nonzero ideal I of R a positive integer $|I|$ such that $|I| = 1 \iff I = R$ and $|IJ| = |I||J|$ for all I and J. (Such functions are not mysterious: they are all gotten by assigning to each nonzero prime ideal an integer greater than 1 and extending by multiplicativity.) For instance on $\mathbb{Z}$ we take the usual absolute value. On $k[t]$, we take $|x| = e^{\deg x}$.

Given such an ideal norm, we may define the **covolume** of a lattice $\Lambda \subset R^n$ as

$$\mathrm{Covol}(R^n/\Lambda) = |\chi(R^n/\Lambda)|.$$

If as above $\Lambda = CR^n$ for some $C \in M_n(R) \cap \mathrm{GL}_n(K)$, then we get

$$\mathrm{Covol}(R^n/\Lambda) = |\det C|.$$

This is most of the content of the proof of Tornheim's Linear Forms Theorem.

Corollary 18.4. *Let $n \in \mathbb{Z}^+$, $M \in R$ with $\deg M \geq 1$, $\theta_1, \ldots, \theta_n \in K$. Then there are $\ell_1, \ldots, \ell_n \in R$, $m \in R^\bullet$ such that*
- $\deg m < \deg M$ *and*
- *For $1 \leq i \leq n$, $\deg(m\theta_i - \ell_i) \leq \frac{-\deg M}{n}$.*

Proof. Let

$$C = \begin{bmatrix} -1 & 0 & \ldots & 0 & \theta_1 \\ 0 & -1 & \ldots & 0 & \theta_2 \\ \vdots & & & & \vdots \\ 0 & 0 & \ldots & -1 & \theta_n \\ 0 & 0 & \ldots & 0 & 1 \end{bmatrix}.$$

Then $\det C = (-1)^n$, so $\deg \det C = 1$. Take

$$\epsilon_1 = \ldots = \epsilon_n = \lfloor \frac{-\deg M}{n} \rfloor,$$
$$\epsilon_{n+1} = \deg M - 1.$$

We have

$$n + 1 + \sum_{i=1}^{n+1} \epsilon_i = n + 1 + (\deg M) - 1 + n \lfloor \frac{-degM}{n} \rfloor$$
$$\geq n + \deg M + n \left(\frac{-\deg M}{n} - 1 + \frac{1}{n} \right) = 1 > 0,$$

so Tornheim's Linear Forms Theorem applies: there is $x \in (R^{n+1})^\bullet$ such that $\deg x_{n+1} < \deg M$ and for all $1 \leq i \leq n$, $\deg(\theta x_{n+1} - x_i) \leq \frac{-degM}{n}$. Put $m = x_{n+1}$ and, for $1 \leq i \leq n$, $\ell_i = x_i$.

It remains to check that $x_{n+1} \neq 0$. If $x_{n+1} = 0$, then $\deg(-x_i) \leq \frac{-degM}{n} < 0$, and since $x_i \in R$ this would force $x_1 = \ldots = x_n = 0$ and thus $x = 0$, contradiction. □

18.3. **Eichler's Linear Forms Theorem.**

We continue to denote by k an arbitary field, $R = k[t]$ and $K = k(t)$. Further, we put $K_\infty = k((\frac{1}{t}))$, the completion of K with respect to the discrete valuation $v_\infty(f) = -\deg f$. In this section we wish to present a profound strengthening of Tornheim's linear forms theorem due to M. Eichler.

Theorem 18.5. *(Eichler's Linear Forms Theorem) Let $M = (m_{ij}) \in \mathrm{GL}_n(K_\infty)$, let $M^* = (m_{ij}^*) = (M^t)^{-1}$, and let $e_1, \ldots, e_n \in \mathbb{Z}$. For $1 \leq i \leq n$, put $e_i^* = -2 - e_i$. Denote by V and V^* the finite-dimensional k-vector spaces consisting of solutions to the systems*

$$\deg \sum_{j=1}^n m_{ij} x_j \leq e_i, \ \forall 1 \leq i \leq n \tag{85}$$

and

$$\deg \sum_{j=1}^n m_{ij}^* x_j^* \leq e_i^*, \ \forall 1 \leq i \leq n. \tag{86}$$

a) We have

$$\dim_k V - \dim_k V^* = \sum_{i=1}^n (e_i + 1) - \deg \det M.$$

b) For all $v = (x_1, \ldots, x_n) \in V$ and $v^* = (x_1^*, \ldots, x_n^*) \in V^*$, we have

$$\sum_{i=1}^n x_i x_i^* = 0.$$

Proof. If $M \in \mathrm{GL}_n(K_\infty)$, we put $M^* = (M^t)^{-1}$. Also put $\deg M = \min \deg m_{ij}$: or, really, $\deg M = \max(-\deg m_{ij})$.
a) Step 1: We assume that $M \in \mathrm{GL}_n(K)$. ...
Step 2: We assume that the result holds for all $M \in \mathrm{GL}_n(K)$ and show that it holds for all $M \in \mathrm{GL}_n(K_\infty)$. Set

$$y_i^* = \sum_j m_{ij}^* x_j,$$

so that

$$x_j^* = \sum_i y_i^* m_{ij}.$$

Conisder instead of (86) the equivalent system

$$\deg y_i^* \leq -2 - e_j, \ \sum_i y_i^* m_{ij} \in k[x]. \tag{87}$$

We appoximate each m_{ij} by a Laurent polynomial:

$$m_{ij} = m_{ij}^{(v)} + r_{ij}^{(v)}, -\deg r_{ij}^{(v)} \geq v.$$

Thus, in matrix notation, $M = M^{(v)} + R^{(v)}$. We apply the matrix identity

$$M^{-1} - N^{-1} = M^{-1}(N - M)N^{-1}$$

to get

$$M^* = M^{(v)*} - M^{(v)*}(R^{(v)})^t M^* = M^* + S^{(v)},$$

say. We have $\lim_{v\to\infty} \deg S^{(v)} = \lim_{v\to\infty} \deg(R^{(v)})^t = -\infty$, so $\lim_{v\to\infty} M^{(v)*} = M^*$. Let $V^{(v)}$ and $V^{(v)*}$ be the dimensions of the solution spaces to (85) and (87) with $M^{(v)}$ and $M^{(v)*}$ in place of M and M^*. Now observe: • Since $\deg\det$: $\mathrm{GL}_n(K_\infty \to \mathbb{Z}$ is continuous, it is eventually constant, ,and thus $\deg\det M^{(v)} = \deg\det M$ for all sufficiently large v.
• For sufficiently large v we have $V^v = V$ and $V^{(v)*} = V^*$. (We leave this to the reader for now.) Thus from the truth of the theorem for all matrices $M \in \mathrm{GL}_n(K)$ the result follows.
b) Because M^* is the inverse transpose of M, we have

$$\sum_{i=1}^n x_i x_i^* = x^t I_n x^* = x^t (M^*)^t M x = (M^* x^*)^t M x.$$

Since $v \in V$, $v^* \in V^*$, it follows that

$$\deg \sum_{i=1}^n x_i x_i^* = \deg((M^* x^*)^t M x) \leq \max_i (e_i - 2 - e_i) = -2.$$

But only the zero polynomial has negative degree, so $\sum_{i=1}^n x_i x_i^* = 0$. □

Exercise: Deduce Tornheim's Linear Forms Theorem from Theorem 18.5.

18.4. **Function Field Vinogradov Lemma.**

Theorem 18.6. *Let k be a field, and let $R = k[t]$. Let $a, b, n \in R^\bullet$ with $\deg n \geq 1$ and $\gcd(ab, n) = 1$. Let $e_1, e_2 \in \mathbb{N}$ be such that*

$$e_1 + e_2 \geq \deg n - 1. \tag{88}$$

Then there are $x, y \in R$, not both zero, such that
(i) $ax \equiv by \pmod n$, and
(ii) $\deg x \leq e_1$, $\deg y \leq e_2$.

Proof. Since $\gcd(b, n) = 1$, there exists $a' \in R$ with $ba' \equiv a \pmod n$. Now consider the linear system with defining matrix $C = \begin{bmatrix} a' & n \\ 1 & 0 \end{bmatrix}$. Note that $\deg\det C = \deg n$. Thus if $e_1, e_2 \in \mathbb{N}$ satisfy (88), then we have

$$\deg\det C = \deg n = (\deg n - 1) + 1 \leq e_1 + e_2 + 1 < e_1 + e_2 + 2,$$

so by Theorem 18.3 there are $X, Y \in R$, not both zero, such that

$$\deg L_1(X,Y) = \deg a'X + nY \leq e_1,$$

$$\deg L_2(X,Y) = \deg X \leq e_2.$$

Put $x = L_1(X,Y) = a'X + nY$, $y = L_2(X,Y) = X$, so $(x,y) \in (R^2)^\bullet$, $\deg x \leq e_1$, $\deg y \leq e_2$ and $ax = a(\frac{b}{a}X + nY) \equiv bX \equiv by \pmod{n}$. □

Remark: The result is trivial when $\deg n = 0$: for then $n \in R^\times$ and the congruence (i) holds for all x and y. Thus $\deg n > 0$ is analogous to the requirement $n > 1$ in Vinogradov's Lemma over $\mathbb{Z}$.

Exercise: When $\deg n = 1$, Theorem 18.6 says that we may take x and y to be constant polynomials. Prove this directly.

18.5. **Prestel's Isotropy Theorem.**

Let k be a field, of characteristic different from 2 but otherwise aribtrary. Put $R = k[t]$ and $K = k(t)$. Fix a positive integer $q \geq 2$. (When k is finite, in some sense the best choice is $q = \#k$, but the choice of q will be absolutely immaterial for the results of this section.) We define a function $|\cdot| : R \to \mathbb{N}$, by $f \in R \mapsto |f| = q^{\deg f}$. Our convention is that the zero polynomial has degree $-\infty$ and that $q^{-\infty} = 0$, i.e., $|0| = 0$. It is immediate to see that $|\cdot|$ satisfies the following properties:

(N1) For $x \in R$, $|x| = 0 \iff x = 0$.
(N2) For $x \in R$, $|x| = 1 \iff x \in R^\times$.
(N3) For $x, y \in R$, $|xy| = |x||y|$.

We extend $|\cdot|$ to a function from K to $\mathbb{Q}^{\geq 0}$ by multiplicativity:

$$|\frac{x}{y}| = \frac{|x|}{|y|}.$$

Lemma 18.7. *The norm function defines an ultrametric on K. In particular, for all $x, y \in K$ we have $|x + y| \leq \max |x|, |y|$.*

Proof. Exercise. □

For $v = (x_1, \ldots, x_n) \in K^n$, we put

$$|v| = \max_i |x_i|.$$

Similarly, for a matrix $A = (a_{ij}) \in M_n(K)$, we put

$$||A|| = \sum_{i,j} |a_{ij}|.$$

Finally, for an n-ary quadratic form q with coefficients in K, we put

$$||q|| = ||A||,$$

where $A \in M_n(K)$ is the corresponding symmetric matrix, i.e., $q(t) = t^T A t$.

By an **isotropic vector** v for a quadratic form q, we mean a nonzero vector

$v \in R^n$ such that $q(v) = 0$. We say that q is **isotropic** if it has an isotropic vector; otherwise we say q is **anisotropic**.

Theorem 18.8. *(Prestel* [Pr87]*) Let $q = q(t_1, \ldots, t_n)$ be an n-ary quadratic form with coefficients in R, $q \not\equiv 0$. If q is isotropic, there is an isotropic vector v with*

$$0 \leq ||v|| \leq ||q||^{\frac{n-1}{2}}.$$

Proof. Let $a = (a_1, \ldots, a_n) \in R^n$ be a nonzero anisotropic vector for q with $||a||$ minimal. By relabelling the variables if necessary, we may assume $||a|| = |a_1|$. Further, we may assume $|a_1| \geq 2$: if $||a|| = |a_1| = 1$, then $||a|| = 1 \leq (||q||^{\frac{n-1}{2}}$. For $v, w \in K^n$, we define the associated bilinear form

$$\langle v, w \rangle = \frac{q(v+w) - q(v) - q(w)}{2}.$$

For all $v \in K^n$, $\langle v, v \rangle = q(v)$.
Step 1: We claim that for all $b \in K^n$ with $q(b) \neq 0$ and all $c \in K$,

$$||q||^{\frac{-1}{2}} \leq ||b - ca||.$$

Let

$$a^* = q(b)a - 2\langle a, b \rangle b.$$

Then $a^* = q(b)\tau_b(a)$, where $\tau_b \in O(q)$ is reflection through the anisotropic vector b. It follows that since $q(a) = 0$, $q(a^*) = 0$. Or, if you like, just calculate directly:

$$\begin{aligned} q(a^*) = \langle a^*, a^* \rangle &= \langle q(b)a - 2\langle a, b \rangle b, q(b)a - 2\langle a, b \rangle b \rangle \\ &= q(b)^2 \langle a, a \rangle - 4\langle a, b \rangle^2 q(b) + 4\langle a, b \rangle^2 q(b) = 0, \end{aligned}$$

since $q(a) = 0$. By the minimality of a, we have

$$||a|| \leq ||a^*||. \tag{89}$$

Now put $d = b - ca$, so $b = d + ca$. Then

$$\begin{aligned} a^* &= q(d + ca)a - 2\langle a, d + ca \rangle (d + ca) \\ &= (q(d) + 2c\langle a, d \rangle + c^2 q(a))a - 2\langle a, d(d + ca) \\ &= q(d)a - 2\langle a, d \rangle d. \end{aligned}$$

Using the ultrametric property of the norm on K, we get

$$||a^*|| = ||q(d)a - 2\langle a, d \rangle d|| \leq ||q||||a||||d||^2. \tag{90}$$

Combining (89) and (90) and dividing through by $||a||$, we get

$$1 \leq ||q|| \cdot ||b - ca||^2,$$

or equivalently

$$||q||^{\frac{-1}{2}} \leq ||b - ca||. \tag{91}$$

Step 2: We claim there is $b \in R^n$ and $c \in K$ with $q(b) \neq 0$ and

$$||b - ca|| \leq ||a||^{\frac{-1}{n-1}}. \tag{92}$$

Apply Corollary with $n - 1$ in place of n, $M = a_1$, $\theta_i = \frac{a_i}{a_1}$. Then there is $0 < b_1 < |a_1|$ and $b_2, \ldots, b_n \in \mathbb{Z}$ such that $|b_i - \frac{b_1}{a_1} a_i| \leq ||a||^{\frac{-1}{n-1}}$. If we take $c = \frac{b_1}{a_1}$ this defines $b \in \mathbb{Z}^n$ with $||b - ca|| \leq ||a||^{\frac{-1}{n-1}}$. Further, for all $2 \leq i \leq n$, we have

$$|b_i| \leq |\frac{b_1}{a_1} a_i| + |a_1|^{\frac{-1}{n-1}} < |b_1| + 1,$$

so $||b|| = |b_1| < |a_1| = ||a||$. By minimality of $||a||$, this forces $q(b) \neq 0$.
Step 3: Combining (91) and (92) we get

$$||q||^{\frac{-1}{2}} \leq ||a||^{\frac{-1}{n-1}},$$

or

$$||a|| \leq ||q||^{\frac{n-1}{2}}.$$

□

Of course our proof of Prestel's Isotropy Theorem is a cut and paste of the proof of Cassels' Isotropy Theorem with some minor modifications (mostly coming from the simpler inequalities afforded by an ultrametric norm). Indeed Prestel explicitly models his proof after that of Cassels. But Prestel's proof is presented slightly differently and – we found – more cleanly than that of Cassels, so indeed the proof we presented of the Cassels Isotropy Theorem is more immediately inspired by Prestel than Cassels.

Also our proof of Step 2 is different from Prestel. Prestel first remarks that the result can be proven using the non-Archimedean GoN of Mahler and Eichler. We find that to be an intriguing remark, but since we have not developed that subject here we have used Tornheim's Linear Forms Theorem instead.

18.6. **The Prospect of a GoN Proof for Ternary Hasse-Minkowski.**

In this section we explore a test case of our "function field geometry of numbers", namely a function field analogue of the Cochrane-Mitchell Theorem.

For the moment, let K be any global field of characteristic different from 2, i.e., a finite extension of either $\mathbb{Q}$ or $\mathbb{F}_q(t)$ for some odd prime power q. We will need (and use without further comment here) the notion of *places* v of K: we denote the set of all such places by Σ_K.

Let $q = q(x_1, \ldots, x_n)$ be a quadratic form over K. For each $v \in \Sigma_K$ we may consider the form $q_v = q_{/K_v}$ over the completion of K at v. Simply because K_v/K is a field extension, it is clear that if q is isotropic, then q_v is isotropic for all v. What is remarkable is that the converse also holds.

Theorem 18.9. *(Hasse-Minkowski) For a quadratic form $q_{/K}$, TFAE:*
(i) q is isotropic: there is $x \in (K^n)^\bullet$ with $q(x) = 0$.
(ii) For all $v \in \Sigma_K$, there is $x_v \in (K_v^n)^\bullet$ with $q_v(x_v) = 0$.

When $n = 3$ there is an extra relation between global and local forms.

Theorem 18.10. *(Hilbert Reciprocity) Let q be a ternary quadratic form over K.*
a) The places v of K such that q_v is anisotropic form a finite set of even cardinality.
b) Let $S \subset \Sigma_K$ be a finite subset. Then there is a ternary quadratic form q over K such that for all $v \in \Sigma_K$, q_v is anisotropic iff $v \in S$.

Remark: When $K = \mathbb{Q}$, Theorem 18.10 is equivalent to the classical quadratic reciprocity law together with the supplementary laws for $(\frac{-1}{p})$ and $(\frac{2}{p})$. Thus Theorem 18.10 can be viewed as a reasonable analogue of quadratic reciprocity in an arbitrary global field. Nowadays it is perhaps best viewed as a statement about quaternion algebras over a global field; we do not intend to discuss these matters here.

Corollary 18.11. *Let q be a ternary form over a global field K. Let v_0 be a place of K. Suppose that for all $v \neq v_0$, q_v is isotropic. Then q is isotropic over K.*

Let us denote by $\mathrm{HM}(K,n)$ the assertion of Hasse-Minkowski for the global field K and for quadratic forms in n variables. The point is that Hasse-Minkowski is in general a difficult theorem – the proof requires deep facts of the arithmetic of global fields coming fom **class field theory** – and the standard proof is not constructive or quantitative. It is a worthy goal to give more elementary, more constructive and more quantitative proofs of the Hasse-Minkowski Theorem in various cases.

Indeed, Legendre's Theorem is an explicit form of $\mathrm{HM}(\mathbb{Q},3)$. Namely, any ternary quadratic form over $\mathbb{Q}$ can be diagonalized and then by elementary reductions reduced to the case of **Legendre Form**

$$q = ax^2 + by^2 + cz^2 = 0, a, b, c \in \mathbb{Z}^\bullet, \ abc \text{ squarefree.}$$

Lemma 18.12. *For a Legendre Form $q_{/\mathbb{Z}}$, consider the following conditions:*
(i) q is isotropic (over $\mathbb{Z}$, or equivalently, over $\mathbb{Q}$).
*(ii) q satisfies the **Legendre conditions**:*
- *$-ab$ is a square modulo c,*
- *$-ac$ is a square modulo b,*
- *$-bc$ is a square modulo a.*

(iii) q_p is isotropic for all odd primes p.
Then (i) $\implies$ (ii) $\iff$ (iii).

The proof of this uses standard techniques (in particular, Hensel's Lemma) and is left to the reader. However, we wish to emphasize the following point: by a CRT argument it is enough to consider congruence conditions modulo p for all primes $p \mid abc$. However, when $p = 2$ asking for an integer to be a square modulo p is vacuous, so the Legendre conditions do not yield any information on the isotropy of $q_{/\mathbb{Q}_2}$. Neither are we getting any conditions on isotropy at $\mathbb{Q}_\infty = \mathbb{R}$, of course. Comparing with Theorem 18.10 we see that the Legendre Conditions are insufficient to force isotropy: we need to include a condition that forces isotropy *either* at 2 or at ∞. In general the arithmetic of global fields at dyadic places (i.e., those for which the residue field has cardinality 2) is unrewardingly complex, whereas the condition for isotropy at an Archimedean place is very simple: we simply need a, b, c to be neither all positive nor all negative. And indeed, Legendre's Theorem says that the Legendre conditions plus the (obviously necessary, equally well in the 18th century as today) sign conditions at ∞ are sufficient for q to be isotropic.

Now let K be any global field, and let R be a PID with fraction field K. Then maximal ideals of R (which correspond to nonzero prime elements of R up to associates) give rise to places of K. To emphasize this, we write Σ_R for the set of maximal ideals of R (in somewhat more generality it is a good idea to let Σ_R denote the *height one prime ideals* of R, but never mind that for now). Let us consider the prospect of proving $\mathrm{HM}(K,3)$ via a Legendre Theorem over R. As usual, given a ternary form $q_{/K}$ we may diagonalize it and clear denominators to get

$$q(x,y,z) = ax^2 + by^2 + cz^2 = 0, \ abc \in R^\bullet;$$

further, using the fact that R is a PID, we reduce to **Legendre Form**, the case in which abc is squarefree.[23] Now we have an analogue of Lemma 18.12:

Lemma 18.13. *For a Legendre Form $q_{/R}$, consider the following conditions:*
(i) q is isotropic (over R, or equivalently, over K).
*(ii) q satisfies the **Legendre conditions**:*
- *$-ab$ is a square modulo c,*
- *$-ac$ is a square modulo b,*
- *$-bc$ is a square modulo a.*

(iii) q_v is isotropic for all nondyadic places $v \in \Sigma_R$.
Then (i) $\implies$ (ii) $\iff$ (iii).

It is natural to supplement the Legendre conditions with sign conditions at the real Archimedean places of K (if any). When are these enough to imply isotropy of q?

Suppose first that K is a number field. Then we must have

$$\mathbb{Z}_K \subset R \subset K.$$

The number field K always has at least one dyadic place, and if there are at least two dyadic places – i.e., if there is more than one prime of $\mathbb{Z}_K$ lying above 2 – our conditions are insufficient. Further, if $\mathbb{Z}_K \subsetneq R$ then there will be non-Archimedean places $v \in \Sigma_K \backslash \Sigma_R$; if there is at least one non-dyadic such place, then our conditions are insufficient. Thus (by using Hasse-Minkowski!) we deduce:

Lemma 18.14. *Let K be a number field. Suppose that there is a unique prime ideal $\mathfrak{p}$ of $\mathbb{Z}_K$ lying over 2, and suppose that $\mathbb{Z}_K[\frac{1}{\mathfrak{p}}]$ is a PID.*[24] *Then a Legendre Form q is isotropic over K iff it satisfies the Legendre conditions and the sign conditions at the real Archimedean places of K, if any.*

Suppose now that K is a finite extension of $\mathbb{F}_q(t)$ for an odd prime power q. In this case there are neither Archimedean places nor dyadic places, so there is the prospect for the Legendre conditions alone to force isotropy. Here though a different phenomenon arises: there is no one choice of R such that every non-Archimedean $v \in \Sigma_K$ comes from a maximal ideal of R; this amounts to the fact that R is an affine coordinate ring and K is the function field of a projective curve which must have at least one closed point "at infinity".

18.7. Chonoles's Geometry of Numbers in $\mathbb{F}_q((\frac{1}{t}))$.

In this section we largely follow the 2012 Honors Thesis of Zev Chonoles [Ch12].

Let K be a field which is locally compact and not discrete. The locally compact abelian group $(K, +)$ admits a **Haar measure**, unique up to scaling. For any automorphism f of $(K, +)$, $S \mapsto \mu(f(S))$ is again a Haar measure on K, so there is a unique scalar $|f| \in \mathbb{R}^{>0}$ such that for all measurable subsets $S \subset K$ we have

$$\mu(f(S)) = |f|\mu(S).$$

[23] Note that this is the only place in which we use that R is a PID. It is thus not clear that this is a crucial restriction.

[24] In general it is not possible to invert a prime ideal. To make sense of it here we use the basic number theoretic fact that there is some $d \in \mathbb{Z}^+$ such that $\mathfrak{p}^d = (x)$ is principal. Then $\mathbb{Z}_K[\frac{1}{x}]$ kills the prime ideal $\mathfrak{p}$ and no other primes: this is what we really mean by $\mathbb{Z}_K[\frac{1}{\mathfrak{p}}]$.

Note that $|f|$ is independent of the chosen Haar measure μ. In particular, for $x \in K^\times$, multiplication by x is an automorphism of K and this defines a real number $|x|$. It turns out that $x \mapsto |x|$ is a norm on K see e.g. [W, §I.2]. Further, using this norm function one can show that K is either $\mathbb{R}$, $\mathbb{C}$, or the norm $x \mapsto |x|$ is of the form $c^{-v(x)}$ for a discrete valuation v on K and some constant $c > 1$ [W, §I.3].

Exercise: a) If $K = \mathbb{R}$, show that $|x|$ is the usual absolute value on $\mathbb{R}$.
b) If $K = \mathbb{C}$, show that $|x|$ is the square of the usual absolute value on $\mathbb{C}$.
c) Suppose $x \mapsto |x|$ is the norm attached to a discrete valuation v on K, with valuation ring R and finite residue field k. Show that $|x| = (\#k)^{-v(x)}$.
d) In particular, if $K = \mathbb{F}_q((\frac{1}{t}))$, show that $|x| = q^{\deg x}$.

Fix $n \in \mathbb{Z}^+$ and put $V = K^n$. Using the above constructed norm on K we introduce a norm on V, $|(x_1, \ldots, x_n)| = \max_i |x_i|$. This norm endows V with the structure of a locally compact topological group. Let μ be a Haar measure on V. As above, for any matrix $M \in \mathrm{GL}(V)$ there is a unique positive real number $|M|$ such that for all measurable subsets $S \subset V$ we have

$$\mu(MS) = |M|\mu(S).$$

Proposition 18.15. *(*[W, p. 7]*) With notation above, we have* $|M| = |\det M|$.

Proof. One reduces to verifying the result for each of the three types of elementary matrices. This makes a good exercise. It is merely alluded to in [W] but done in detail in [Ch12] in the (entirely representative) case $K = \mathbb{F}_q((t))$. □

Let q be a prime power, $A = \mathbb{F}_q[t]$, and $K = \mathbb{F}_q(t)$. Let v_∞ be the discrete valuation on K given by $v_\infty(f/g) = \deg g - \deg f$, and let $|\cdot|_v$ be defined by $|h|_\infty = q^{-v_\infty(h)}$. Let $K_\infty = \mathbb{F}_q((\frac{1}{t}))$ be the completion of K at the place v_∞. This is a locally compact, complete, discretely valued field with valuation ring $R_\infty = \mathbb{F}_q[[\frac{1}{t}]]$, maximal ideal $\mathfrak{m}_\infty = \frac{1}{t}R_\infty$ and finite residue field $\mathbb{F}_q$. Fix $n \in \mathbb{Z}^+$ and let $V = K_\infty^n$ be an n-dimensional vector space over K_∞, with a standard basis $(e_1, \ldots, e_n)$. Thus the additive group of V is a locally compact abelian group so admits a Haar measure, and indeed a unique Haar measure μ such that $\mu(\mathfrak{m}_\infty^n) = 1$. (The most standard normalization would be to give unit mass to R_∞^n, not to m_∞^n. Thus this Haar measure is q^n times the "most standard" one. It will soon become clear why Chonoles's normalization is a good one for geometry of numbers considerations.)

Let $V = K_\infty^n$, and let $e_1, \ldots, e_n$ denote the standard basis of V. An **A-lattice** in V is the A-span of a K_∞-basis of V. Since A is a PID and V is torsionfree, any A-lattice is isomorphic as an A-module to A^n.

Any $x \in K_\infty$ can be uniquely written in the form $f + g$ for $f \in A$ and $g \in \mathfrak{m}_\infty$. Thus for a lattice $\Lambda = \langle v_1, \ldots, v_n \rangle_A$, any $v \in V$ can be uniquely expressed as

$$v = \sum_{j=1}^n f_j v_j + \sum_{j=1}^n g_j v_j, \ f_j \in A, \ g_j \in \mathfrak{m}_\infty.$$

It follows that if we put $D_\Lambda = \sum_{j=1}^n \mathfrak{m}_\infty v_j$, then

$$V = \Lambda \oplus D_\Lambda.$$

Thus D_Λ is a **fundamental domain** for Λ in V which is moreover an additive subgroup (and even an R_∞-module). Accordingly, we define

$$\operatorname{Covol}\Lambda = \mu(D_\Lambda).$$

Let $\mathcal{E} = \sum_{i=1}^n Ae_i$ be the "standard A-lattice" in V. Then $D_\mathcal{E} = \mathfrak{m}_\infty^n$, so by our normalization of the Haar measure we have $\operatorname{Covol}\mathcal{E} = \mu(D_\mathcal{E}) = 1$.

Proposition 18.16. *Let $V = \sum_{i=1}^n Av_i$ be an A-lattice in V, write $v_i = \sum_{j=1}^n m_{ij}e_j$, and put $M = (m_{ij}) \in \operatorname{GL} V$. Then*

$$\operatorname{Covol}\Lambda = |\det M|_\infty.$$

Proof. We have $\Lambda = M\mathcal{E}$ and thus $D_\Lambda = MD_\mathcal{E}$. By Proposition 18.15,

$$\operatorname{Covol}\Lambda = \mu(D_\Lambda) = \mu(MD_\mathcal{E}) = |\det M|_\infty \operatorname{Covol}\mathcal{E} = |\det M|_\infty.$$

□

Theorem 18.17. *Let Λ be an A-lattice in V, and let $\mathcal{B} \subset V$ be a measurable subset which is* closed under subtraction *and satisfies*

$$\mu(\mathcal{B}) > \operatorname{Covol}\Lambda.$$

Then $\mathcal{B} \cap \Lambda^\bullet \neq \varnothing$.

Proof. Since $V = \coprod_{\lambda\in\Lambda} D_\Lambda + \lambda$, we have

$$\mathcal{B} = \coprod_{\lambda\in\Lambda} \left((D_\Lambda + \lambda) \cap \mathcal{B}\right).$$

By countable additivity – note $\Lambda \cong \mathbb{F}_q[t]^n$ is countable! – and translation invariance,

$$\mu(\mathfrak{B}) = \sum_{\lambda\in\Lambda} \mu((D_\Lambda + \lambda) \cap \mathfrak{B}) = \mu((D_\Lambda \cap (\mathfrak{B} - \lambda)) + \lambda) = \mu(D_\Lambda \cap (\mathfrak{B} - \lambda)).$$

If $\{D_\Lambda \cap (\mathfrak{B} - \lambda)\}_{\lambda\in\Lambda}$ were pairwise disjoint, we'd have

$$\mu(\mathfrak{B}) = \sum_{\lambda\in\Lambda} \mu(D_\Lambda \cap (\mathfrak{B} - \lambda)) \leq \mu(D_\Lambda) = \operatorname{Covol}\Lambda,$$

contradicting our hypothesis. Thus there are $c_1, c_2 \in \mathfrak{B}$ and $\lambda_1 \neq \lambda_2 \in \Lambda$ such that

$$c_1 - \lambda_1 = c_2 - \lambda_2 \in D_\Lambda.$$

Since $\mathfrak{B}$ is closed under subtraction, we deduce

$$c_1 - c_2 = \lambda_1 - \lambda_2 \in \mathfrak{B} \cap \Lambda^\bullet.$$

□

We can easily deduce a new proof of Tornheim's Linear Form in the case $k = \mathbb{F}_q$. In fact the argument works for matrices with coefficients in $\mathbb{F}_q((\frac{1}{t}))$, not just $\mathbb{F}_q[t]$.

Corollary 18.18. *(Finite Field Tornheim Theorem) Let q be a prime power; let $C = (c_{ij}) \in M_n(\mathbb{F}_q((\frac{1}{t})))$ with $\det C \neq 0$. Let $e_1, \ldots, e_n \in \mathbb{Z}$ be such that*

$$\deg\det C < \sum_{i=1}^n (e_i + 1). \tag{93}$$

Then there exists $x \in (\mathbb{F}_q[t]^n)^\bullet$ *such that for all* $1 \le i \le n$,

$$\deg\left(\sum_{j=1}^{n} c_{ij}x_j\right) \le e_i. \tag{94}$$

Proof. As above, let $\mathcal{E} = \mathbb{F}_q[t]^n$ be the standard A-lattice in $V = K_\infty^n = \mathbb{F}_q((\frac{1}{t}))^n$. Let $\mathfrak{B}$ be the set of $x \in V$ satisfying (94). We are trying to show that $\mathfrak{B} \cap \mathcal{E}^\bullet \neq \varnothing$. Since $\mathfrak{B}$ is closed under subtraction – indeed, it is an R_∞-submodule of V – by Theorem 18.17 it suffices to show that $\mu(\mathfrak{B}) > \operatorname{Covol}\mathcal{E} = 1$. To see this, consider

$$\mathfrak{b} = \{x = (x_1, \ldots, x_n) \in V \mid \deg x_i \le e_i \ \forall 1 \le i \le n\}.$$

The subset $\mathfrak{b}$ is a Cartesian product of subsets $\mathfrak{b}_i = \{x \in K_\infty \mid v_\infty(x) \ge e_i\}$ of K_∞; thus $\mu(\mathfrak{b}_i) = q^{e_i+1}$ and

$$\mu(\mathfrak{b}) = q^{\sum_{i=1}^n (e_i+1)}.$$

Further, $\mathfrak{B} = C^{-1}\mathfrak{b}$, so

$$\mu(\mathfrak{B}) = |\det C|^{-1}\mu(\mathfrak{b}) = q^{\sum_{i=1}^n (e_i+1) - \deg\det C}.$$

Thus by our hypothesis (93) we have $\mu(\mathfrak{B}) > 1 = \operatorname{Covol}\mathcal{E}$, and the result follows. □

18.8. **Mahler's non-Archimedean Geometry of Numbers.**

Let K be a field, and let $|\cdot| : K \to \mathbb{R}^{\ge 0}$ be a non-Archimedean norm on K:

- $\forall x \in K,\ |x| = 0 \iff |x| = 0.$
- $\forall x, y \in K,\ |xy| = |x||y| \forall x, y \in K.$
- $\forall x, y \in K,\ |x + y| \le \max|x|, |y|.$

Let $\widehat{K}$ be the completion of $(K, |\cdot|)$; let R and $\widehat{R}$ be the corresponding valuation rings. For any $N \in \mathbb{Z}^+$, we may view $\widehat{K}^N$ as a non-Archimedean normed $\widehat{K}$-space by setting, for $x = (x_1, \ldots, x_N) \in \widehat{K}^N$,

$$|x| = \max_i |x_i|.$$

We also define the "standard" (possibly isotropic) inner product on $\widehat{K}^N$:

$$(x_1, \ldots, x_N) \cdot (y_1, \ldots, y_N) = \sum_{i=1}^{N} x_i y_i.$$

All of the following are immediate:

$$\forall x \in \widehat{K}^N, |x| \ge 0,\ |x| = 0 \iff x = 0.$$
$$\forall \alpha \in \widehat{K},\ x \in \widehat{K}^N,\ |\alpha x| = |\alpha||x|.$$
$$\forall x, y \in \widehat{K}^N,\ |x + y| \le \max(|x|, |y|).$$
$$\forall x, y \in \widehat{K}^N,\ |x \cdot y| \le |x||y|.$$

Now, as we did in Euclidean space, consider certain properties of $f : \widehat{K}^N \to \mathbb{R}$:

(DF1′): $f(0) = 0$.
(DF1): $\forall x \in \widehat{K}^N,\ f(x) = 0 \iff x = 0$.

(DF2) $\forall \alpha \in \widehat{K},\ x \in \widehat{K}^N,\ f(\alpha x) = |\alpha| f(x)$.
(DF3) $\forall x, y \in \widehat{K}^N,\ f(x+y) \leq \max(f(x), f(y))$.

A function f satisfying (DF1′), (DF2) and (DF3) is a **pseudo-distance function**, and a function f satisfying (DF1), (DF2) and (DF3) is a **distance function**.

For a pseudo-distance function f and $\tau \in \mathbb{R}^{>0}$, we put

$$\Omega_{f,\tau} = f^{-1}([0,\tau]).$$

By definition, a subset $\Omega \subset \widehat{K}^n$ is **convex** if it is of the form $\Omega_{f,\tau}$ for some pseudo-distance function f and some $\tau > 0$. If we may take f to be a distance function, we say Ω is a **convex body**.

Condition (DF2) is analogous to the symmetry condition $f(-x) = f(x)$ in the classical case, but is much stronger. The following exercise drives this point home.

Exercise: Show that any convex subset $\Omega \subset \widehat{K}^N$ is a $\widehat{R}$-submodule of $\widehat{K}^N$.

Theorem 18.19. *Let $f : \widehat{K}^N \to \mathbb{R}$ be a pseudo-distance function.*
a) If $C_f = \max_{1 \leq i \leq n} f(e_i)$, then for all $x \in \widehat{K}^N$ we have

$$f(x) \leq C_f |x|.$$

b) If f is a distance function, there exists $\gamma_f \in \mathbb{R}^{>0}$ such that for all $x \in \widehat{K}^N$,

$$c_f |x| \leq f(x).$$

Corollary 18.20. *For a convex subset $\Omega \subset \widehat{K}^N$, the following are equivalent:*
(i) Ω is a convex body.
(ii) Ω is bounded.

Proof. (i) $\implies$ (ii): Let Ω be a convex body, so $\Omega = f^{-1}([0,R])$ for a distance function f. Then for $x \in \Omega$, $c_f|x| \leq f(x) \leq R$, so $|x| \leq \frac{R}{c_f}$, so Ω is bounded.
$\neg$ (i) $\implies$ $\neg$ (ii): If Ω is a convex set but not a convex body, then $\Omega = f^{-1}([0,R])$ for a pseudodistance function f which is not a distance function, i.e., for which there is $0 \neq x \in \widehat{K}^N$ such that $f(x) = 0$. Then by homogeneity $f(\alpha x) = 0$ for all $\alpha \in \widehat{K}$, and thus Ω contains the entire line $\langle x \rangle_{\hat{K}}$ so is unbounded. □

Exercise: Let $f_1, f_2 : \widehat{K}^N \to \mathbb{R}$ be two pseudodistance functions. Show that $\max(f_1, f_2)$ is a pseudodistance function.

Exercise: Let $\{\Omega_i\}_{i \in I}$ an indexed family of convex subsets of $\widehat{K}^N$.
a) Show that $\Omega = \bigcap_{i \in I}$ is a convex set.
b) Show that if at least one Ω_i is a convex body, then Ω is a convex body.

As in the Archimedean case, the easiest pseudodistance functions come from *linear forms*: if $L : \widehat{K}^N \to \widehat{K}$ is any linear form, then $|L|$ is a pseudodistance function on $\widehat{K}^N$, which is not a distance function except in the trivial case $N = 1$. However, if we are given a system $L_1, \ldots, L_N$ of linear forms – say $L_i(x) = \sum_{j=1}^N \alpha_{ij} x_j$ – then by the exercises above $P = \max_{i=1}^N |L_i|$ is a pseudodistance function, which (by nothing else than elementary linear algebra) is a distance function iff the the

corresponding matrix $M = (\alpha_{ij})$ is invertible, and the associated convex bodies $P^{-1}([0, R])$ will be called **parallelopipeds**.

18.9. **Normed Rings.**

A **norm** on a ring R is a function $|\cdot| : R \to \mathbb{N}$ such that
(N0) $|x| = 0 \iff x = 0$,
(N1) $\forall x, y \in R$, $|xy| = |x||y|$, and
(N2) $\forall x \in R$, $|x| = 1 \iff x \in R^\times$.

A **normed ring** is a pair $(R, |\cdot|)$ where $|\cdot|$ is a norm on R. A nonzero ring admitting a norm is necessarily a domain. We denote the fraction field by K.

Let R be a domain with fraction field K. We say that two norms $|\cdot|_1, \cdot|\cdot|_2$ on R are **equivalent** – and write $|\cdot|_1 \sim |\cdot|_2$ if for all $x \in K$, $|x|_1 < 1 \iff |x|_2 < 1$.

Remark 3.1: Let $(R, |\cdot|)$ be a normed domain with fraction field K. By (N1) and (N2), $|\cdot| : (R^\bullet, \cdot) \to (\mathbb{Z}^+, \cdot)$ is a homomorphism of commutative monoids. It therefore extends uniquely to a homomorphism on the group completions, i.e., $|\cdot| : K^\times \to \mathbb{Q}^{>0}$ via $|\frac{x}{y}| = \frac{|x|}{|y|}$. This map factors through the **group of divisibility** $G(R) = K^\times/R^\times$ to give a map $K^\times/R^\times \to \mathbb{Q}^{>0}$.

Example 3.2: The usual absolute value $|\cdot|_\infty$ on $\mathbb{Z}$ (inherited from $\mathbb{R}$) is a norm.

Example 3.3: Let k be a field, $R = k[t]$, and let $a \geq 2$ be an integer. Then the map $f \in k[t]^\bullet \mapsto a^{\deg f}$ is a non-Archimedean norm $|\cdot|_a$ on R and the norms obtained for various choices of a are equivalent. As we shall see, when k is finite, the most natural normalization is $a = \#k$. Otherwise, we may as well take $a = 2$.

Example 3.4: Let R be a discrete valuation ring (DVR) with valuation $v : K^\times \to \mathbb{Z}$ and residue field k. For any integer $a \geq 2$, we may define a norm on R, $|\cdot|_a : R^\bullet \to \mathbb{Z}^{>0}$ by $x \mapsto a^{v(x)}$. (Note that these are the *reciprocals* of the norms $x \mapsto a^{-v(x)}$ attached to R in valuation theory.) Using the fact that $G(R) = K^\times/R^\times \cong (\mathbb{Z}, +)$ one sees that these are all the norms on R. That is, a DVR admits a unique norm up to equivalence.

Example 3.5: Let R be a UFD. Then $\mathrm{Prin}(R)$ is a free commutative monoid on the set Σ_R of height one primes of R. Thus, to give a norm map on R it is necessary and sufficient to map each prime element π to an integer $n_\pi \geq 2$ in such a way that if $(\pi) = (\pi')$, $n_\pi = n_{\pi'}$.

A norm $|\cdot|$ on a ring R is **metric** if for all $x, y \in R$, $|x + y| \leq |x| + |y|$. A norm is **ultrametric** if for all $x, y \in R$, $|x + y| \leq \max |x|, |y|$.

Example 18.21. *The standard norm (Euclidean absolute value) on $\mathbb{Z}$ is metric.*

Example 18.22. *For any field k and any $a \geq 2$, on the ring $R = k[t]$ the norm $f \in R \mapsto a^{\deg f}$ is ultrametric: indeed, for $f, g \in R$,*

$$|f + g| = a^{\deg(f+g)} \leq a^{\max \deg f, \deg g} = \max a^{\deg f}, a^{\deg g} = \max |f|, |g|.$$

Example 18.23. *Let R be a discrete valuation ring which is not a field, with valuation $v : R^\bullet \to \mathbb{Z}$, $v(0) = -\infty$. Then for any $a \geq 2$ and $x \in R$, putting $|x| = a^{v(x)}$ gives a norm on R. But* ***beware****: this norm is* not *ultrametric nor even metric. Indeed, let π be a uniformizing element $x = \pi^2 - 1$, $y = 1$. Then*

$$a^{v(x+y)} = a^{v(\pi^2)} = a^2 > 1 + 1 = a^{v(\pi^2-1)} + a^{v(1)}.$$

Notice in particular that our definition of the norm attached to a discrete valuation is the reciprocal *of the usual definition, and thus the metric properties are lost.*

In fact among all normed rings, examples of metric norms – and still more, ultrametric norms – seem to be quite rare. We get a slightly larger class of examples by relaxing the metric condition, as follows.
Let $|\cdot|$ be a norm on a ring R. Define

$$A(R) = \inf\{A \in \mathbb{R}^{>0} \mid \forall x, y \in R, |x + y| \leq A(|x| + |y|)\},$$
$$C(R) = \inf\{A \in \mathbb{R}^{>0} \mid \forall x, y \in R, |x + y| \leq C \max |x|, |y|\}.$$

The following result connects some simple facts about these quantities.

Lemma 18.24. *a) If $A(R) < \infty$, then for all $x, y \in K$, $|x + y| \leq A(R)(|x| + |y|)$.*
b) If $C(R) < \infty$, then for all $x, y \in K$, $|x + y| \leq C(R) \max |x|, |y|)$.
c) We have $A(R) \leq C(R) \leq 2A(R)$.
d) In particular, $A(R) < \infty \iff C(R) < \infty$.

Proof. Exercise. □

We call a norm **almost metric** if $A(R) < \infty$ (equivalently by Lemma 18.24, if $C(R) < \infty$). Note that a norm is metric if $A(R) \leq 2$ and ultrametric iff $C(R) = 1$.

Theorem 18.25. *Let K_0 denote either $\mathbb{Q}$ or $\mathbb{F}_p(t)$. Let K/K_0 be a finite separable extension of degree d. Let S be a finite, nonempty set of places of K containing all Archimedean places (if any), and let R be the ring of S-integers of K. TFAE:*
(i) $\#S = 1$.
(ii) The unit group $R^\times$ is finite.
(iii) The canonical norm function $x \in R^\bullet \mapsto \#R/(x)$ is almost metric.
(iv) $C(R) = 2^d$.

Proof. ... □

Corollary 18.26. *Let R be an S-integer ring in a number field K.*
a) The canonical norm on R is almost metric iff
(i) $K = \mathbb{Q}$ and $R = \mathbb{Z}$, or
(ii) K is imaginary quadratic and $R = \mathbb{Z}_K$ is the full ring of integers.
b) In case (i) above, $C = 2$. In case (ii) above, $C = 4$.

For a normed Dedekind domain $(R, |\cdot|)$, we define the **Euclideanity**

$$E(R) = \sup_{x \in K} \inf_{y \in R} |x - y|.$$

As usual, we say that $|\cdot|$ is a **Euclidean norm** on R if for all $x \in K$ there exists $y \in R$ with $|x - y| < 1$. Thus in particular R is Euclidean if $E(R) < 1$ and is not Euclidean if $E(R) > 1$. Because of the supremum in the definition of $E(R)$, the case $E(R) = 1$ is ambiguous: *a priori* it is possible for a ring with $E(R) = 1$ to be Euclidean, but in every example I know with $E(R) = 1$, the norm is *not* Euclidean. In any case, we really will want to use the stronger condition $E(R) < 1$ in our work

below, so this distinction is not really relevant for us.

As is well-known, in a Euclidean ring every ideal is generated by each element of minimal norm, so a Euclidean ring is a PID.

Example 18.27. *Let $R = \mathbb{Z}$ endowed with the standard absolute value. Then $E(R) = \frac{1}{2}$, so $\mathbb{Z}$ is Euclidean.*

Example 18.28. *Let k be any field, $R = k[t]$, and let $a \geq 2$ be an integer. Endow R with the nom $|f|_a = a^{\deg f}$. Then $E(R) = \frac{1}{a}$, so R is Euclidean.*

Note that we may have $E(R) = \infty$; we say R is **E-finite** if $E(R) < \infty$.

Lemma 18.29. *Let R be a PID with fraction field K, and let $|\cdot|$ be a metric norm on R. Let L/K be a finite separable field extension, and let S be the integral closure of R in L, endowed with its extended norm. Then S is an E-finite Dedekind domain.*

Proof. Let $n = [L : K]$. It is a standard result in algebraic number theory that S is a Dedekind domain (this does not use the hypothesis of separability) and that $S \cong R^n$ (this does!). Let $\sigma_1, \ldots, \sigma_n : L \hookrightarrow \overline{K}$ be the n-distinct K-algebra embeddings into an algebraic closure, so for $x \in L$, $|x| = |\prod_{i=1}^n \sigma_i(x)|$. Let $x_1, \ldots, x_n$ be an R-basis for S, hence also a K-basis for L. Therefore, for any $x \in L$, there are unique $\alpha_1, \ldots, \alpha_n \in L$ such that $x = \sum_i \alpha_i x_i$. Fix $\epsilon > 0$, and choose for all i an element $\beta_i \in R$ such that $|\alpha_i - \beta_i| \leq E(R) + \epsilon$. Then

$$|x - \sum_{i=1}^n \beta_i x_i| \leq \sum_{i=1}^n |\alpha_i - \beta_i||x_i| \leq (E(R) + \epsilon) \sum_{i=1}^n |x_i|.$$

Thus S is E-finite. □

18.10. **Gerstein-Quebbemann.**

The following result is an abstraction of Hermite's proof of Theorem 10.1.

Theorem 18.30. *Let $(R, |\cdot|)$ be an almost metric normed ring with $E(R) < 1$.*
a) Suppose $A(R)E(R)^2 < 1$. Then for all $n \geq 2$,

$$\gamma_n(R) \leq \left(\frac{A(R)}{1 - A(R)E(R)^2}\right)^{\frac{n-1}{2}}.$$

b) Suppose $C(R)E(R)^2 < 1$. Then for all $n \geq 2$,

$$\gamma_n(R) \leq C(R)^{\frac{n-1}{2}}.$$

c) If R is ultrametric, $\gamma_n(R) \leq 1$ for all $n \in \mathbb{Z}^+$.

Proof. The greater part of the argument involves deriving the inequality (95) below. Combining this with $|x+y| \leq A(R)(|x| + |y|)$ we deduce part a); combining it with $|x + y| \leq C(R) \max |x|, |y|$, we deduce part b).

Since $E(R) < 1$, R is Euclidean and thus a PID, so Hermite's Lemma applies. Let $q = \sum_{ij} a_{ij} t_i t_j : K^n \to K$ be an anisotropic quadratic form. By Hermite's Lemma, after making a unimodular change of variables we may assume that the minimum of q on R^n is attained at the first standard basis vector e_1.

Let $\varphi : K^n \to K^n$ be the K-linear map given by $e_1 \mapsto e_1' = e_1$, $e_j \mapsto e_j' = e_j - \frac{a_{1j}}{a_{11}}$

for $2 \leq j \leq n$, so e_1 is orthogonal to the subspace $\langle e_2', \ldots, e_n' \rangle$. Note also that $\det \varphi = 1$. Let

$$q'(t) = q(\varphi(t)) = a_{11}t_1^2 + q_2(t_2, \ldots, t_n).$$

Then $\operatorname{disc} q_2 = \frac{\operatorname{disc} q'}{a_{11}} = \frac{\operatorname{disc} q}{a_{11}}$. Now for $\lambda_1, \ldots, \lambda_n \in R$, write

$$w = (\lambda_1 + \frac{a_{12}}{a_{11}}\lambda_2 + \ldots + \frac{a_{1n}}{a_{11}}\lambda_n)e_1 + \lambda_2 e_2' + \ldots + \lambda_n e_n' = \gamma e_1 + z,$$

say. Suppose z is chosen so as to be minimal for q_2 on $\bigoplus_{i=2}^n Re_i'$. Then

$$|q(z)| = |q_2(\lambda_2, \ldots, \lambda_n)| \leq \gamma_{n-1}(R)|\operatorname{disc} q_2|^{\frac{1}{n-1}} = \gamma_{n-1}(R)|a_{11}|^{\frac{-1}{n-1}}|\operatorname{disc} q|^{\frac{1}{n-1}}.$$

Let $\epsilon > 0$ be small enough so that $E(R) + \epsilon < 1$. By definition of $E(R)$, there is $\lambda_1 \in R$ with $|\gamma| \leq E(R) + \epsilon < 1$. Thus we have

$$\min(q) = |a_{11}| \leq |q(w)| = |\gamma^2 a_{11} + q_2(\lambda_2, \ldots, \lambda_n)|. \tag{95}$$

a) By definition of $A(R)$, we have

$$|a_{11}| \leq |\gamma^2 a_{11} + q_2(\lambda_2, \ldots, \lambda_n)| \leq A(R)\left(|\gamma^2||a_{11}| + |q_2(\lambda_2, \ldots, \lambda_n)|\right)$$

$$\leq A(R)(E(R) + \epsilon)^2|a_{11}| + A(R)\gamma_{n-1}(R)|a_{11}|^{\frac{-1}{n-1}}|\operatorname{disc} q|^{\frac{1}{n-1}}.$$

Since this inequality holds for all sufficiently small ϵ, it also holds for $\epsilon = 0$:

$$|a_{11}| \leq A(R)E(R)^2|a_{11}| + A(R)\gamma_{n-1}(R)|a_{11}|^{\frac{-1}{n-1}}|\operatorname{disc} q|^{\frac{1}{n-1}}.$$

Multiplying through by $|a_{11}|^{\frac{1}{n-1}}$ gives

$$|a_{11}|^{\frac{n}{n-1}} \leq A(R)E(R)^2|a_{11}|^{\frac{n}{n-1}} + A(R)\gamma_{n-1}(R)|\operatorname{disc} q|^{\frac{1}{n-1}},$$

and thus

$$\frac{|a_{11}|^n}{|\operatorname{disc} q|} \leq \left(\frac{A(R)}{1 - A(R)E(R)^2}\right)^{n-1} \gamma_{n-1}(R)^{n-1}.$$

This implies

$$\gamma_n(R)^n \leq \left(\frac{A(R)}{1 - A(R)E(R)^2}\right)^{n-1} \gamma_{n-1}(R)^{n-1}$$

and thus

$$\gamma_n(R) \leq \left(\frac{A(R)}{1 - A(R)E(R)^2}\right)^{\frac{n-1}{n}} \gamma_{n-1}(R)^{\frac{n-1}{n}}.$$

Using $\gamma_1(R) = 1$, an easy induction argument gives

$$\gamma_n(R) \leq \left(\frac{1}{1 - A(R)E(R)^2}\right)^{\frac{n-1}{2}},$$

completing the proof of part a). As for part b), starting again from (95) we get

$$|a_{11}| \leq = |\gamma^2 a_{11} + q_2(\lambda_2, \ldots, \lambda_n)| \leq C(R)\max(|\gamma|^2|a_{11}|, |q_2(\lambda_2, \ldots, \lambda_n)|)$$

and thus (inserting and then removing an $\epsilon > 0$ as above) we get

$$|a_{11}| \leq C(R)\max(E(R)^2|a_{11}|, \gamma_{n-1}(R)|a_{11}|^{\frac{-1}{n-1}}|\operatorname{disc} q|^{\frac{1}{n-1}}).$$

But by our hypothesis, $|a_{11}| > C(R)E(R)^2|a_{11}|$, so we must have

$$|a_{11}| \leq C(R)\gamma_{n-1}(R)|a_{11}|^{\frac{-1}{n-1}}|\operatorname{disc} q|^{\frac{1}{n-1}}).$$

Thus

$$|a_{11}|^{\frac{n}{n-1}} \leq C(R)\gamma_{n-1}(R)|\operatorname{disc} q|^{\frac{1}{n-1}}$$

and hence

$$\frac{|a_{11}|^n}{|\operatorname{disc} q|} \leq C^{n-1}\gamma_{n-1}^{n-1}(R).$$

Taking nth roots gives

$$\gamma_n(R) \leq C^{\frac{n-1}{n}}\gamma_{n-1}^{\frac{n-1}{n}}.$$

Exactly as in part a), an easy induction argument gives $\gamma_n(R) \leq C^{\frac{n-1}{2}}$.
c) Since $|\cdot|$ is ultrametric iff $C(R) = 1$, this follows immediately from part b). □

Exercise: Check that Theorem 18.30 implies Theorem 10.1.

Corollary 18.31. *Let $R = \mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ be the ring of integers of the imaginary quadratic field $K = \mathbb{Q}(\sqrt{-3})$. Then for all $n \in \mathbb{Z}^+$, we have*

$$\gamma_n(R) \leq \left(\frac{36}{5}\right)^{\frac{n-1}{2}}.$$

Proof. For the ring of integers $\mathbb{Z}_K$ of an imaginary quadratic field K we have

$$E(\mathbb{Z}_K) = \frac{|m|+1}{4},\ Z_K = \mathbb{Z}[\sqrt{-m}],$$

$$E(\mathbb{Z}_K) = \frac{(|m|+1)^2}{16m},\ \mathbb{Z}_K = \mathbb{Z}[\frac{1+\sqrt{-m}}{2}].$$

By Lemma 18.24 and Theorem 18.25, $A(R) \leq C(R) = 4$. Since $A(R)E(R)^2 < 1$, Theorem 18.30a) applies. □

Remark: Rather disappointingly, it turns out that $R = \mathbb{Z}$ and $R = \mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ are the only two S-integer rings to which the hypotheses of Theorem 18.30 apply!

Corollary 18.32. *(Gerstein* [Ge73]*, Quebbemann) Let k be a field of characteristic different from 2. Let $a \geq 2$ be an integer. Endow $R = k[t]$ with the norm $|f|_a = a^{\deg f}$. Then $\gamma_n(k[t]) \leq 1$ for all $n \in \mathbb{Z}^+$.*

Proof. Since $E(R) = \frac{1}{a}$, $C(R) = 1$, $C(R)E(\mathbb{R})^2 < 1$, and Theorem 18.30c) applies. □

Remark 4.11: A field k of characteristic not 2 admits an ordering iff for all $n \in \mathbb{Z}^+$ the quadratic form $q_n = \langle 1, \ldots, 1\rangle = t_1^2 + \ldots + t_n^2$ is anisotropic. Such a field k is necessarily infinite, and then an easy specialization argument shows that any anisotropic form q over k remains anisotropic upon base extension to $k[t]$. So let k be a field admitting an ordering – e.g. $k = \mathbb{R}$ or any of its subfields. Then for all $n \in \mathbb{Z}^+$, $|\operatorname{disc} q_n| = 1$ and $m(q_n) = 1$, so $\gamma(q_n) = 1$ and thus $\gamma_n(k[t]) = 1$. This shows that the bound of Corollary 18.32 is best possible without further restrictions on k. On the other hand, if $k = \mathbb{F}_q$ then every quadratic form in at least 5 variables over $k[t]$ is isotropic, and thus $\gamma_n(\mathbb{F}_q[t]) = 0$ for all $n \geq 5$.

Theorem 18.33. *(Samuel* [**?**]*) Suppose that X is a curve of genus zero. Then $S = k[X]$ is Euclidean with respect to the extended norm iff the projective closure $\overline{X}$ of X is isomorphic to $\mathbb{P}^1$ and the gcd of the degrees of the points at infinity is 1.*

19. Abstract Blichfeldt and Minkowski

We wish to develop an "abstract" version of Blichfeldt's Lemma. This begins with a **measured group** $(G, +, \mathcal{A}, \mu)$: a group $(G, +)$ – not assumed to be commutative, though we write the group law additively – and a measure $(G, \mathcal{A}, \mu)$ which is right invariant: for all $A \in \mathcal{A}$ and $x \in G$, $\mu(A + x) = \mu(A)$. To avoid trivialities, we assume $\mu(G) > 0$.

Let Γ be a subgroup of G. A **fundamental domain** $\mathcal{F}$ for Γ in G is a measurable subset $\mathcal{F} \subset G$ such that
(FD1) $\bigcup_{g \in \Gamma} \mathcal{F} + g = \Gamma$, and
(FD2) Fo all $g_1, g_2 \in \Gamma$, $\mu((\mathcal{F} + g_1) \cap (\mathcal{F} + g_2)) = 0$.

Lemma 19.1. *If $\mathcal{F}_1$ and $\mathcal{F}_2$ are both fundamental domains for a countable subgroup Γ in G, then $\mu(\mathcal{F}_1) = \mu(\mathcal{F}_2)$.*

Proof. Observe that if $\{\mathcal{S}_i\}_{i \in I}$ is a countable family of subsets such that $\mu(S_i \cap S_j) = 0$ for all $i \neq j$, then

$$\mu(\bigcup_{i \in I} S_i) = \sum_{i \in I} \mu(S_i).$$

Now we have

$$\mathcal{F}_1 \supset \mathcal{F}_1 \cap (\bigcup_{g \in \Gamma} \mathcal{F}_2 + g) = \bigcup_{g \in \Gamma} \mu(\mathcal{F}_1 \cap (\mathcal{F}_2) + g),$$

so, using the above observation,

$$\mu(\mathcal{F}_1) \geq \sum_{h \in H} \mu(\mathcal{F}_1 \cap (\mathcal{F}_2 + g)) = \sum_{g \in \Gamma} \mu(\mathcal{F}_1 \cap (\mathcal{F}_2 - g)) = \sum_{g \in \Gamma} \mu((\mathcal{F}_1 + g) \cap \mathcal{F}_2)$$

$$= \mu(\bigcup_{g \in \Gamma} (\mathcal{F}_1 + g) \cap \mathcal{F}_2) = \mu(\mathcal{F}_2).$$

Interchanging $\mathcal{F}_1$ and $\mathcal{F}_2$ we get the result. □

Example: Let G be a Lie group, and let Γ be a discrete subgroup of G. Then Γ is countable, and there is a fundamental domain $\mathcal{F}$ for Γ in G, which can moreover be taken to be regular-closed, i.e., equal to the closure of its interior.

We say a subgroup Λ of a measured group G is a **lattice** if it is countable and admits a measurable fundamental domain of finite measure. We define the **covolume** $\operatorname{Covol} \Lambda$ to be the measure of any such fundamental domain.

Exercise: Show that for a lattice Γ in a measured group G, $\operatorname{Covol} \Lambda > 0$. (Hint: recall our assumption that $\mu(G) > 0$.)

Theorem 19.2. *(Abstract Blichfeldt Lemma) Let Λ be a lattice in a measured group G, and let $M \in \mathbb{Z}^+$. Let $\Omega \subset G$ be measurable, and suppose*

$$\frac{\mu(\Omega)}{\operatorname{Covol} \Lambda} > M. \tag{96}$$

There are distinct $w_1, \ldots, w_{M+1} \in \Omega$ such that for all $1 \leq i, j \leq M+1$, $w_i - w_j \in \Lambda$.

Proof. Let $\mathcal{F}$ be a measurable fundamental domain for Λ in G. For $x \in \Lambda$, let

$$\Omega_x = \Omega \cap (\mathcal{F} + x).$$

Then $\Omega = \bigcup_{x \in \Gamma} \Omega_x$: this is a countable union which is essentially pairwise disjoint – for all $x \neq y \in \Gamma$, $\mu(\Omega_x \cap \Omega_y) = 0$ – so

$$\sum_{x \in \Gamma} \mu(\Omega_x - x) = \sum_{x \in \Lambda} \mu(\Omega_x) = \mu(\Omega) > M \operatorname{Covol}(\Lambda) = M\mu(\mathcal{F}). \tag{97}$$

We apply the Measure Theoretic Pigeonhole Principle with $X = \mathcal{F}$, $I = \Lambda$, $S_x = \Omega_x - x$: there is $v \in \mathcal{F}$ and $x_1, \ldots, x_{M+1} \in \Lambda$ such that

$$v \in \bigcap_{i=1}^{M+1} \Omega_{x_i} - x_i.$$

Thus for $1 \leq i \leq M+1$ there is $w_i \in \Omega_{x_i}$ – so $w_1, \ldots, w_{M+1}$ are distinct – with

$$\forall 1 \leq i \leq M+1, w_i - x_i = v.$$

It follows that for all $1 \leq i, j \leq M+1$, $w_i - w_j = (x_i + v) - (x_j + v) = x_i - x_j \in \Lambda$. □

A **measured ring** is a ring endowed with a measure such that the additive group of R is a measured group. Again we assume $\mu(R) > 0$ to avoid trivialities.

Theorem 19.3. *(Abstract Minkowski Theorem) Let $M \in \mathbb{Z}^+$, $(R, +, \cdot, \mathcal{A}, \mu)$ be a measured ring, and let $\Lambda \subset R^N$ be a countable subgroup. Let $\Omega \subset R$ be measurable and symmetric: $x \in \Omega \implies -x \in \Omega$.*
a) We suppose $2 \in R^\bullet$ and all of the following:
- *Ω is **midpoint closed**: $x, y \in \Omega \implies \frac{x+y}{2} \in \Omega$.*
- *2Λ is a lattice in R.*
- *$\frac{\mu(\Omega)}{\operatorname{Covol} 2\Lambda} > M$.*

Then $\#(\Omega \cap \Lambda^\bullet) \geq M$.
b) We suppose all of the following:
- *Ω is closed under subtraction: $x, y \in \Omega \implies x - y \in \Omega$.*
- *Λ is a lattice in R.*
- *$\frac{\mu(\Omega)}{(\operatorname{Covol} \Lambda} > M$.*

Then $\#(\Omega \cap \Lambda^\bullet) \geq M$.

Proof. a) Apply the Abstract Blichfeldt Lemma with $G = (R, +)$ and 2Λ in place of Λ. We get distinct elements $w_1, \ldots, w_{M+1} \in \Omega$ such that for all $1 \leq i, j \leq M+1$, $\frac{w_i - w_j}{2} \in \Lambda$. Since Ω is symmetric and midpoint closed, $-w_j \in \Omega$ and thus $\frac{w_i - w_j}{2} \in \Omega$ for all $1 \leq i, j \leq M+1$. Fixing $i = 1$ and letting j run from 2 to $M+1$ gives us M nonzero elements of $\Omega \cap \Lambda$.
b) This is exactly the same as part a) except we use Λ instead of 2Λ and use the fact that Ω is closed under subtraction. □

Remark: Suppose R is a locally compact topological ring, μ is a Haar measure on $(R, +)$ and $\Lambda \subset R$ is a discrete subring. Then in every "natural example" I know Λ will necessarily be countable and there will be a measurable set of coset representatives for Λ in R. Because Haar measures are Radon measures and hence finite on compact subsets, Λ will be a lattice if it admits a compact fundamental domain. Note also that our fundamental Blichfeldt conditions depend only on the ratio of the volume of Ω and the volume of a fundamental domain for 2Λ or Λ, hence is independent of the choice of Haar measure.

Example: Take $R = \mathbb{R}^N$ and $\Lambda \subset \mathbb{R}^N$ a (full) lattice in the usual sense: we recover Minkowski's Convex Body Theorem.

Example: Let q be a prime power. Take $R = \mathbb{F}_q((\frac{1}{t}))^N$ and Λ a $\mathbb{F}_q[t]$-lattice in R. We recover Chonoles's Convex Body Theorem.

References

[A2.5] P.L. Clark, *Algebra Handout 2.5: More on Commutative Groups*. Lecture notes available at `http://www.math.uga.edu/∼pete/4400algebra2point5.pdf` Euclidean quadratic forms and ADC forms I. (pdf), Acta Arithmetica 154 (2012), 137-159

[AC05] G. Alon and P.L. Clark, *On the Number of Representations of an Integer by a Linear Form*. Journal of Integer Sequences, Vol. 8 (2005), Article 05.5.2.

[ADCI] P.L. Clark, *Euclidean quadratic forms and ADC forms I.* Acta Arithmetica 154 (2012), 137–159.

[ADCII] P.L. Clark and W.C. Jagy, *Euclidean quadratic forms and ADC-Forms II: integral forms*. Acta Arith. 164 (2014), no. 3, 265-308.

[And67] G.E. Andrews, *Classroom Notes: On the Geometry of Numbers in Elementary Number Theory*. Amer. Math. Monthly 74 (1967), 1124-1125.

[Ank57] N.C. Ankeny, *Sums of three squares*. Proc. Amer. Math. Soc. 8 (1957), 316-319.

[Ba08] R. Bacher, *A new inequality for the Hermite constants*. Int. J. Number Theory 4 (2008), 363-386.

[BWZ65] R.P. Bambah, A. Woods and H. Zassenhaus, *Three proofs of Minkowski's second inequality in the geometry of numbers*. J. Austral. Math. Soc. 5 (1965), 453-462.

[Bh00] M. Bhargava, *On the Conway-Schneeberger fifteen theorem*. Quadratic forms and their applications (Dublin, 1999), 2737, Contemp. Math., 272, Amer. Math. Soc., Providence, RI, 2000.

[BD58] B.J. Birch and H. Davenport, *Quadratic equations in several variables*. Proc. Cambridge Philos. Soc. 54 (1958), 135-138.

[BH] M. Bhargava and J.P. Hanke, *Universal quadratic forms and the 290-Theorem*. Preprint.

[BSD56] B.J. Birch and H.P.F. Swinnerton-Dyer, *On the inhomogeneous minimum of the product of n linear forms*. Mathematika 3 (1956), 25-39.

[Bl48] H. Blaney, *Indefinite quadratic forms in n variables*. J. London Math. Soc. 23 (1948), 153-160.

[Bl14] H.F. Blichfeldt, *A new principle in the geometry of numbers, with some applications*. Trans. Amer. Math. Soc. 15 (1914), 227-235.

[Bl19] H.F. Blichfeldt, *Report on the theory of the geometry of numbers*. Bull. Amer. Math. Soc. 25 (1919), 449-453.

[BK05] J. Bochnak and W. Kucharz, *On successive minima of indefinite quadratic forms*. Enseign. Math. (2) 51 (2005), 319-330.

[BO09] J. Bochnak and B.-K. Oh, *Almost-universal quadratic forms: an effective solution of a problem of Ramanujan*. Duke Math. J. 147 (2009), 131-156.

[BR51] A. Brauer and R.L. Reynolds, *On a theorem of Aubry-Thue*. Canadian J. Math. 3 (1951), 367-374.

[Bu] E.B. Burger, *Exploring the number jungle: a journey into Diophantine analysis*. Student Mathematical Library, 8. American Mathematical Society, Providence, RI, 2000.

[Bu96] E.B. Burger, *Small solutions to systems of linear congruences over number fields*. Symposium on Diophantine Problems (Boulder, CO, 1994). Rocky Mountain J. Math. 26 (1996) 875-888.

[C] J.W.S. Cassels, *An introduction to the geometry of numbers*. Corrected reprint of the 1971 edition. Classics in Mathematics. Springer-Verlag, Berlin, 1997.

[CA] P.L. Clark, *Commutative Algebra*. Notes available at `http://math.uga.edu/∼pete/integral.pdf`

[Ca49] J.W.S. Cassels, *The Markoff chain*. Ann. of Math. (2) 50 (1949), 676-685.

[Ca53a] J.W.S. Cassels, *A short proof of the Minkowski-Hlawka theorem.* Proc. Cambridge Philos. Soc. 49 (1953), 165-166.

[Ca53b] J.W.S. Cassels, *Yet another proof of Minkowski's theorem on the product of two inhomogeneous linear forms.* Proc. Cambridge Philos. Soc. 49 (1953), 365-366.

[Ca55] J.W.S. Cassels, *Bounds for the least solutions of homogeneous quadratic equations.* Proc. Cambridge Philos. Soc. 51 (1955), 262-264.

[CF] T.W. Cusick and M.E. Flahive, *The Markoff and Lagrange spectra.* Mathematical Surveys and Monographs, 30. American Mathematical Society, Providence, RI, 1989.

[Ch80] J.H.H. Chalk, *Linearly independent zeros of quadratic forms over number fields.* Monatsh. Math. 90 (1980), 13-25.

[Ch12] Z. Chonoles, *Hermite's Theorem for Function Fields.* Honors Thesis, Brown University, 2012.

[CK09] H. Cohn and A. Kumar, *Optimality and uniqueness of the Leech lattice among lattices.* Ann. of Math. (2) 170 (2009), 1003-1050.

[CLRR80] M.D. Choi, T.Y. Lam, B. Reznick and A. Rosenberg, *Sums of squares in some integral domains.* J. Algebra 65 (1980), 234–256.

[Cl11] P.L. Clark, *Euclidean Quadratic Forms and ADC-forms I.* To appear in *Acta Arith.* Preprint available at `http://math.uga.edu/ pete/ADCForms1.pdf`.

[Cl12] P.L. Clark, *Thue-Vinogradov and Idoneal Quadratic Forms*, preprint.

[Co87] T. Cochrane, *Small solutions of congruences over algebraic number fields.* Illinois J. Math. 31 (1987), 618-625.

[Co89] T. Cochrane, *Small zeros of quadratic forms modulo p.* J. Number Theory 33 (1989), 286-292.

[Co93] T. Cochrane, *On representing the multiple of a number by a quadratic form.* Acta Arith. 63 (1993), 211-222.

[CoMi98] T. Cochrane and P. Mitchell, *Small solutions of the Legendre equation.* J. Number Theory 70 (1998), 62-66.

[Coh] H. Cohen, *A course in computational number theory.* Springer Graduate Texts in Mathematics...

[Col97] L. Colzani, *Approximation of Lebesgue integrals by Riemann sums and lattice points in domains with fractal boundary.* Monatsh. Math. 123 (1997), no. 4, 299-308.

[Con00] J.H. Conway, *Universal quadratic forms and the fifteen theorem.* Quadratic forms and their applications (Dublin, 1999), 23-26, Contemp. Math., 272, Amer. Math. Soc., Providence, RI, 2000.

[CP62] H. Cohn and G. Pall, *Sums of four squares in a quadratic ring.* Trans. Amer. Math. Soc. 105 (1962), 536–552.

[Cr99] J.E. Cremona, *Reduction of binary cubic and quartic forms.* LMS J. Comput. Math. 2 (1999), 64-94.

[CR03] J.E. Cremona and D. Rusin, *Efficient solution of rational conics.* Math. Comp. 72 (2003), 1417-1441.

[Da47] H. Davenport, *The geometry of numbers.* Math. Gaz. 31, (1947), 206-210.

[DaRo47] H. Davenport and C.A. Rogers, *Hlawka's theorem in the geometry of numbers.* Duke Math. J. 14 (1947), 367-375.

[Da52] H. Davenport, *Recent progress in the geometry of numbers.* Proceedings of the International Congress of Mathematicians, Cambridge, Mass., 1950, vol. 1, pp. 166-174. Amer. Math. Soc., Providence, R. I., 1952.

[Da57] H. Davenport, *Note on a theorem of Cassels.* Proc. Cambridge Philos. Soc. 53 (1957), 539-540.

[De02] J.I. Deutsch, *Geometry of numbers proof of Götzky's four-squares theorem.* J. Number Theory 96 (2002), 417-431.

[De04] J.I. Deutsch, *An alternate proof of Cohn's four squares theorem.* J. Number Theory 104 (2004), 263-278.

[De08a] J.I. Deutsch, *Short proofs of the universality of certain diagonal quadratic forms.* Arch. Math. (Basel) 91 (2008), 44-48.

[De08b] J.I. Deutsch, *Bumby's technique and a result of Liouville on a quadratic form.* Integers 8 (2008), no. 2, A2, 20 pp.

[De08c] J.I. Deutsch, *A quaternionic proof of the universality of some quadratic forms.* Integers 8 (2008), no. 2, A3, 23 pp.

[DH48] H. Davenport and M. Hall, *On the equation $ax^2 + by^2 + cz^2 = 0$*. Quart. J. Math., Oxford Ser. 19 (1948), 189-192.

[Dic19] L.E. Dickson, *Applications of the geometry of numbers to algebraic numbers*. Bull. Amer. Math. Soc. 25 (1919), 453-455.

[Dic27] L.E. Dickson, *Integers represented by positive ternary quadratic forms*. Bull. Amer. Math. Soc. 33 (1927), 63-70.

[Die03] R. Dietmann, *Small solutions of quadratic Diophantine equations*. Proc. London Math. Soc. 86 (2003), 545-582.

[DL78] M. Duggal and I.S. Luthar, *Minkowski's theorems in completions of A-fields of non-zero characteristic*. Colloq. Math. 38 (1977/78), no. 2, 329-337.

[Eh55] E. Ehrhrart, *Une généralisation du théorème de Minkowski*. C. R. Acad. Sci. Paris 240 (1955), 483-485.

[Fr33] O. Frink, *Jordan measure and Riemann integration*. Ann. of Math. (2) 46 (1933), 518–526.

[Ga09] É. Gaudron, *Géométrie des nombres adélique et lemmes de Siegel généralisés*. Manuscripta Math. 130 (2009), 159-182.

[Ge73] L.J. Gerstein, *A new proof of a theorem of Cassels and Pfister*. Proc. Amer. Math. Soc. 41 (1973), 327-328.

[Ge79] L.J. Gerstein, *Unimodular quadratic forms over global function fields*. J. Number Theory 11 (1979), 529–541.

[Ge03] L.J. Gerstein, *Definite quadratic forms over $\mathbb{F}_q[t]$*. J. Algebra 268 (2003), 252-263.

[G] L.J. Gerstein, *Basic quadratic forms*. Graduate Studies in Mathematics, 90. American Mathematical Society, Providence, RI, 2008.

[GiPa04] A. Gica and L. Panaitopol, *A result similar to a theorem of Lagrange*. Math. Rep. (Bucur.) 6(56) (2004), 45-50.

[GG06] H. Gillet and D.R. Grayson, *Volumes of symmetric spaces via lattice points*. Doc. Math. 11 (2006), 425-447.

[Gr27] J.H. Grace, *The Four Square Theorem*. J. London Math. Soc. 2 (1927), 3–8.

[Gr07] P.M. Gruber, *Convex and discrete geometry*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 336. Springer, Berlin, 2007.

[Ha11] T. Hagedorn, *Primes of the form $x^2 + ny^2$ and the geometry of (convenient) numbers*, preprint.

[Ha38] P.R. Halmos, *Note on almost-universal forms*. Bull. Amer. Math. Soc. 44 (1938), 141-144.

[Ha64] J. Hammer, *On a general area-perimeter relation for two-dimensional lattices*. Amer. Math. Monthly 71 (1964), 534-535.

[Ha66] J. Hammer, *Some relatives of Minkowski's theorem for two-dimensional lattices*. Amer. Math. Monthly 73 (1966), 744-746.

[Ha68] J. Hammer, *On some analogies to a theorem of Blichfeldt in the geometry of numbers*. Amer. Math. Monthly 75 (1968), 157-160.

[HW6ed] G.H. Hardy and E.M. Wright, *An introduction to the theory of numbers*.

[He53] C. Hermite, Comptes Rendus Paris 37, 1853.

[Hl43] E. Hlawka, *Zur Geometrie der Zahlen*. Math. Z. 49 (1943), 285-312.

[Ho50] L. Holzer, *Minimal solutions of Diophantine equations*. Canadian J. Math. 2 (1950), 238-244.

[HT12] C. Hurlburt and J.L. Thunder, *Hermite's constant for function fields*. Canad. J. Math. 64 (2012), no. 2, 301–317.

[Hu99] W. Hu, *On Minkowski constant of function fields*. Northeast. Math. J. 15 (1999), 69-75.

[Hu00] M.N. Huxley, *Integer points in plane regions and exponential sums*. Number theory, 157-166, Trends Math., Birkhäuser, Basel, 2000.

[Ic97] M.I. Icaza, *Hermite constant and extreme forms for algebraic number fields*. J. London Math. Soc. (2) 55 (1997), 11-22.

[J1] N. Jacobson, *Basic algebra. I*. Second edition. W. H. Freeman and Company, New York, 1985

[JK] W.C. Jagy and I. Kaplansky, *Positive definite binary quadratic forms that represent the same primes*, preprint.

[Jo70] J.-R. Joly, *Sommes des carrés dans certains anneaux principaux*. Bull. Sci. Math. (2) 94 (1970), 85-95.

[Kn59] M. Kneser, *Kleine Lösungen der diophantischen Gleichung* $ax^2 + by^2 = cz^2$. Abh. Math. Sem. Univ. Hamburg 23 (1959), 163-173.

[Ku87] T. Kubota, *Geometry of numbers and class field theory*. Japan. J. Math. (N.S.) 13 (1987), 235-275.

[Le65] W.J. Leahey, *A note on a theorem of I. Niven*. Proc. Amer. Math. Soc. 16 (1965), 1130-1131.

[Mac61] A.M. Macbeath, *Factorization of matrices and Minkowski's conjecture*. Proc. Glasgow Math. Assoc. 5 (1961), 86-89.

[Ma41] K. Mahler, *An analogue to Minkowski's geometry of numbers in a field of series*. Ann. of Math. (2) 42, (1941), 488-522.

[Ma44] K. Mahler, *On a theorem of Minkowski on lattice points in non-convex point sets*. J. London Math. Soc. 19 (1944), 201-205.

[Ma45] K. Mahler, *On lattice points in n-dimensional star bodies I. Existence Theorems*

[Ma46] K. Mahler, *The theorem of Minkowski-Hlawka*. Duke Math. J. 13 (1946), 611-621.

[Ma71] K. Mahler, *A lecture on the geometry of numbers of convex bodies*. Bull. Amer. Math. Soc. 77 (1971), 319-325.

[Ma83] K. Mahler, *On a theorem in the geometry of numbers in a space of Laurent series*. J. Number Theory 17 (1983), 403-416.

[McF71] R.B. McFeat, *Geometry of numbers in adele spaces*. Dissertationes Math. Rozprawy Mat. 88 (1971), 49 pp.

[McM05] C.T. McMullen, *Minkowski's conjecture, well-rounded lattices and topological dimension*. J. Amer. Math. Soc. 18 (2005), 711-734.

[Me09] B. Meyer, *Generalised Hermite constants, Voronoi theory and heights on flag varieties*. Bull. Soc. Math. France 137 (2009), 127-158.

[Mo35] L.J. Mordell, *On some arithmetical results in the geometry of numbers*. Compositio Math. 1 (1935), 248-253.

[Mo44] L.J. Mordell, *Observation on the minimum of a positive quadratic form in eight variables*. J. London Math. Soc. 19 (1944), 3-6.

[Mor48] L.J. Mordell, *The minimum of a definite ternary quadratic form*. J. London Math. Soc. 23 (1948), 175-178.

[Mo51] L.J. Mordell, *On the equation* $ax^2+by^2-cz^2 = 0$. Monatsh. Math. 55 (1951), 323-327.

[Mo53] L.J. Mordell, *Note on Sawyer's paper "The product of two non-homogeneous linear forms."*. J. London Math. Soc. 28 (1953), 510-512.

[Mo58] L.J. Mordell, *On the representation of a number as a sum of three squares*. Rev. Math. Pures Appl. 3 (1958), 25-27.

[Mo66a] L.J. Mordell, *Solvability of the equation* $ax^2+by^2 = p$. J. London Math. Soc. 41 (1966), 517-522.

[Mo66b] L.J. Mordell, *The representation of numbers by some quaternary quadratic forms*. Acta Arith. 12 (1966/1967), 47-54.

[Mo69] L.J. Mordell, *On the magnitude of the integer solutions of the equation* $ax^2+by^2+cz^2 = 0$. J. Number Theory 1 (1969), 1-3.

[MoWa01] M. Morishita and T. Watanabe, *Adèle geometry of numbers*. Class field theoryits centenary and prospect (Tokyo, 1998), 509-536, Adv. Stud. Pure Math., 30, Math. Soc. Japan, Tokyo, 2001.

[MuTh07] M.R. Murty and N. Thain, *Pick's theorem via Minkowski's theorem*. Amer. Math. Monthly 114 (2007), 732-736.

[NCA] P.L. Clark, *Non-commutative algebra*. Lecture notes: `http://math.uga.edu/~pete/noncommutativealgebra.pdf`

[Ni40] I. Niven, *Integers of quadratic fields as sums of squares*. Trans. Amer. Math. Soc. 48 (1940), 405-417.

[NZM] I. Niven, H.S. Zuckerman and H.L. Montgomery, *An introduction to the theory of numbers*. Fifth edition. John Wiley & Sons, Inc., New York, 1991.

[Nu10] L.M. Nunley, *Geometry of Numbers Approach to Small Solutions of the Extended Legendre Equation*. UGA Master's thesis, 2010.

[OhWa01] S. Ohno and T. Watanabe, *Estimates of Hermite constants for algebraic number fields*. Comment. Math. Univ. St. Paul. 50 (2001), 53=63.

[OLD] C.D. Olds, A. Lax and G.P. Davidoff, *The geometry of numbers*. Appendix I by Peter D. Lax. Anneli Lax New Mathematical Library, 41. Mathematical Association of America, Washington, DC, 2000.

[Op46] A. Oppenheim, *Remark on the minimum of quadratic forms*. J. London Math. Soc. 21 (1946), 251–252.

[Pf97] A. Pfister, *Small zeros of quadratic forms over algebraic function fields*. Acta Arith. 79 (1997), 221-238.

[Pr87] A. Prestel, *On the size of zeros of quadratic forms over rational function fields*. J. Reine Angew. Math. 378 (1987), 101-112.

[Ra75] S. Raghavan, *Bounds for minimal solutions of Diophantine equations*. Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II 1975, no. 9, 109-114.

[Ra17] Ramanujan, *On the expression of a number in the form* $ax^2 + by^2 + cz^2 + du^2$, Proceedings of the Cambridge Philosophical Society 19(1917), 11–21. See also `http://en.wikisource.org/wiki/Proceedings_of_the_Cambridge_Philosophical_Society`

[Re70] S. Reich, *Two-dimensional lattices and convex domains*. Math. Mag. 43 (1970), 219-220.

[Ro47] C.A. Rogers, *Existence theorems in the geometry of numbers*. Ann. of Math. (2) 48 (1947), 994-1002.

[Ro51] C.A. Rogers, *The number of lattice points in a star body*. J. London Math. Soc. 26 (1951), 307-310.

[Ro56] C.A. Rogers, *The number of lattice points in a set*. Proc. London Math. Soc. (3) 6 (1956), 305-320.

[Ro64] C.A. Rogers,, *Packing and covering*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 54 Cambridge University Press, New York 1964.

[Ro74] K. Rogers, *Legendre's theorem and quadratic reciprocity*, J. Number Theory 6 (1974), 339-344.

[RoSD58] K. Rogers and H.P.F. Swinnerton-Dyer, *The geometry of numbers over algebraic number fields*. Trans. Amer. Math. Soc. 88 (1958), 227-242.

[S] P. Samuel, *Théorie algébrique des nombres*. Hermann, Paris 1967.

[Sa48] D.B. Sawyer, *The product of two non-homogeneous linear forms*. J. London Math. Soc. 23 (1948), 250-251.

[Sa68] D.B. Sawyer, *Lattice points in rotated star sets*. J. London Math. Soc. 43 (1968), 131-142.

[Sc39] A. Scholz, *Einfuhrung in die Zahlentheorie*, Berlin, 1939.

[Sc93] A. Schrijver, *Graphs on the torus and geometry of numbers*. J. Combin. Theory Ser. B 58 (1993), no. 1, 147-158.

[Si45] C.L. Siegel, *A mean value theorem in geometry of numbers*. Ann. of Math. (2) 46 (1945), 340-347.

[S] C.L. Siegel, *Lectures on the geometry of numbers*. Notes by B. Friedman. Rewritten by Komaravolu Chandrasekharan with the assistance of Rudolf Suter. With a preface by Chandrasekharan. Springer-Verlag, Berlin, 1989.

[Sk52] Th. Skolem, *A simple proof of the condition of solvability of the Diophantine equation* $ax^2 + by^2 + cz^2 = 0$. Norske Vid. Selsk. Forh., Trondheim 24 (1952), 102-107.

[SK68] H. Stevens and L. Kuty, *Applications of an elementary theorem to number theory*. Arch. Math. (Basel) 19 (1968), 37-42.

[SS90] H.P. Schlickewei and W.M. Schmidt, *Bounds for zeros of quadratic forms*. Number theory, Vol. II (Budapest, 1987), 951-964, Colloq. Math. Soc. János Bolyai, 51, North-Holland, Amsterdam, 1990.

[SWO10] K. Sawatani, T. Watanabe and K. Okuda, *A note on the Hermite-Rankin constant*. J. Théor. Nombres Bordeaux 22 (2010), 209-217.

[Th96] J.L. Thunder, *An adelic Minkowski-Hlawka theorem and an application to Siegel's lemma*. J. Reine Angew. Math. 475 (1996), 167-185.

[Th00] J.L. Thunder, *Remarks on adelic geometry of numbers*. Number theory for the millennium, III (Urbana, IL, 2000), 253-259, A K Peters, Natick, MA, 2002.

[To41] L. Tornheim, *Linear forms in function fields*. Bull. Amer. Math. Soc. 47 (1941), 126-127.

[OW00] Z.M. Ou and K.S. Williams, *Small solutions of* $\phi_1 x_1^2 + \cdots + \phi_n x_n^2 = 0$. Canad. J. Math. 52 (2000), no. 3, 613–632.

[Pa51] G. Pall, *Sums of two squares in a quadratic field*. Duke Math. J. 18 (1951), 399-409.
[Va10] S. Vance, *A Mordell inequality for lattices over maximal orders*. Trans. Amer. Math. Soc. 362 (2010), 3827-3839.
[Va11] S. Vance, *Improved sphere packing lower bounds from Hurwitz lattices*. Adv. Math. 227 (2011), 2144-2156.
[Ve13] A. Venkatesh, *A note on sphere packings in high dimension*. Int. Math. Res. Not. IMRN 2013, 1628-1642.
[vHC06] M. van Hoeij and J. Cremona, *Solving conics over function fields*. J. Théor. Nombres Bordeaux 18 (2006), 595-606.
[Vi27] I.M. Vinogradov, *On a general theorem concerning the distribution of the residues and non-residues of powers*. Trans. Amer. Math. Soc. 29 (1927), 209–217.
[W] A. Weil, *Basic Number Theory*.
[Wi71] K.S. Williams, *Note on a theorem of Pall*. Proc. Amer. Math. Soc. 28 (1971), 315-316.
[Wi73] K.S. Williams, *Another proof of a theorem of Niven*. Math. Mag. 46 (1973), 39.
[Wi88] K.S. Williams, *On the Size of a Solution of Legendre's Equation*. Utilitas Math. 34 (1988), 65–72.
[Wi91] J.M. Wills, *Quermassintegrals and related convexity functions in geometry of numbers*. Rend. Sem. Mat. Messina Ser. II 1 (14) (1991), 255-264.
[Wó72] J. Wójcik, *On sums of three squares*. Colloq. Math. 24 (1971/72), 117-119.