Davis Math Lab Project Report, Spring 2026
(Distribution of error terms in lattice point counting)

(Ruihan (Robin) Ma, David Reagon, Yu (Wilson) Wen)
(Adrian Weitzer) (Graduate Mentor)
(Junxian Li) (Faculty Mentor)

June 20, 2026

Contents

1 Introduction

1.1 The Gauss Circle problem

The motivation for our project is a classical problem of number theory: The Gauss Circle problem. First considered by Gauss, it asks: given a circle of radius R centered at α, consider:

$$N_\alpha(R,d) = \#\{n \in \mathbb{Z}^d : |\alpha - n| \le R\}.$$

The set of integer lattice points contained in a circle of radius R, centered at $\alpha \in [0,1]^d$. Asymptotically:

$$N_\alpha(R,d) = \mathrm{Vol}(R,d) + E_\alpha(R,d), \quad R \to \infty.$$

where $\mathrm{Vol}(R,d)$ is the volume of a d-dimensional ball, and $E_\alpha(R,d)$ is the error term. The classical Gauss circle problem concerns pointwise bounds for $E_0(R,d)$ when $d = 2$.

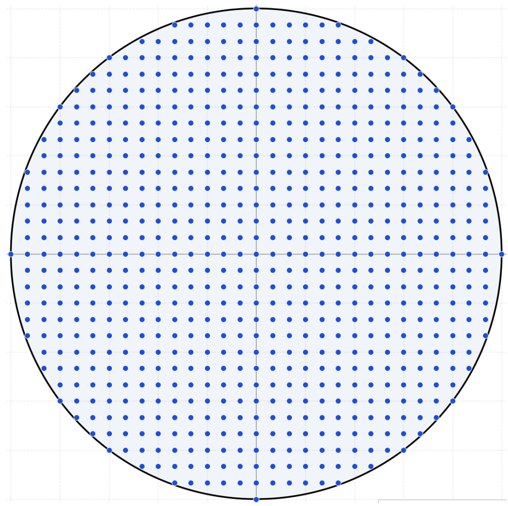


Figure 1: Lattice points in a circle, $R = 15$, $N_0(15, 2) = 709$

In the 1920's Hardy-Littlewood [HL] conjectured that

$$E_0(R, 2) = O(R^{1/2+\varepsilon}) \quad \forall \varepsilon > 0.$$

It can be shown using elementary methods that the error is at most R:

$$E_\alpha(R, 2) = O(R).$$

This conjecture remains open, and the best progress in this direction comes from Huxley [Hux03], who showed:

$$E_0(R, 2) = O(R^{46/73}).$$

1.2 Distribution of Lattice points

Getting pointwise bounds for $E_\alpha(R, 2)$ for a fixed R is quite challenging, requiring nontrivial estimations of exponential sums. Starting in the 90's attention turned towards the distribution of $E_\alpha(R, 2)$ as R varies. To do this, we consider the following normalized error terms:

$$F_\alpha(R, 2) = \frac{N_\alpha(R, 2) - \pi R^2}{R^{1/2}} = \frac{E_\alpha(R, 2)}{R^{1/2}}.$$

[BCDL] showed that this normalized error problem had a limiting distribution as $R \to \infty$, and where $\alpha \in [0, 1]^2$. To do this, he used the following ergodic theorem:

Theorem([BCDL]): If there exists a sequence of periodic functions $\tilde{a}_n$ and coefficients $\tilde{\gamma}_n$ such that:

$$\lim_{N\to\infty} \limsup_{T\to\infty} \frac{1}{T} \int_0^T \min\left\{1, \left| F(t) - \sum_{n\le N} \tilde{a}_n(\tilde{\gamma}_n\, t) \right| \right\} dt = 0.$$

Then $F(t)$ has a limiting distribution as $T \to \infty$.

Furthermore, [BCDL] showed that the limiting distribution could be wildly different depending on the choice of α. (See [B99] and Table 1)

Using this and a formula for the error term in terms of trigonometric functions, [BCDL] was able to show

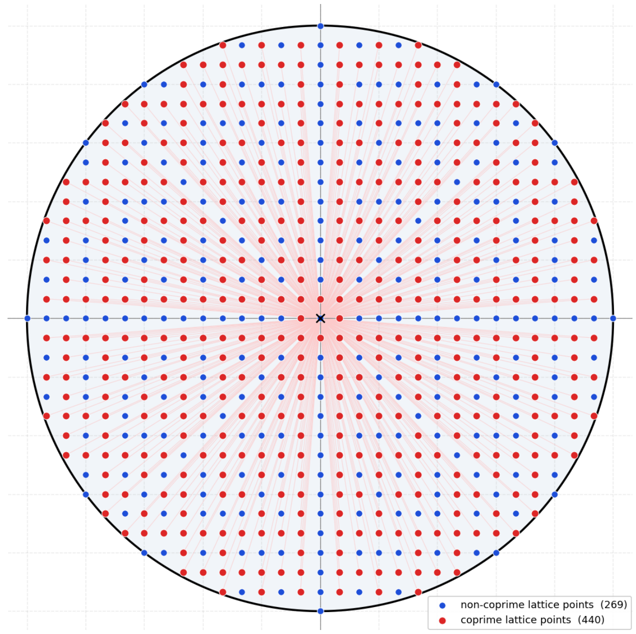


Figure 2: Circle with $R = 15$, Visible lattice points are marked in red.

that a limiting distribution exists and was able to show that it was non gaussian, matching with numerical simulations we ran (see Table 3).

1.3 Primitive Lattice points

A variation that we will consider is that of primitive lattice points, geometrically these are points that are "visible" from the origin (see Figure 2), we detect them using the greatest common divisor, as these are exactly the points whose coordinates are coprime,

$$N^*_\alpha(R, d) = \#\{n \in \mathbb{Z}^d : |\alpha - n| \leq R, \;\; \gcd(n_1, ..., n_d) = 1\}.$$

We denote the primitive lattice point counting function with a star $(*)$ to help differentiate it with the standard lattice point count.

Given the normalized error terms for the higher-dimensional lattice points counting:

$$F_\alpha(R, d) = R^{-\frac{d-1}{2}}\Big(N_\alpha(R, d) - \mathrm{Vol}(R, d)\Big)$$

we have the normalized error terms for the primitive lattice points counting:

$$F^*_\alpha(R, d) = R^{-\frac{d-1}{2}}\left(N^*_\alpha(R, d) - \frac{1}{\zeta(d)}\mathrm{Vol}(R, d)\right),$$

where $\zeta(d)$ is the Riemann-Zeta function evaluated at d. Much of our work this quarter involves understanding the primitive error term, including many numerical simulations.

2 Progress

The theorem mentioned in the last sections requires the error term to be written as a sum of periodic functions. [BCDL] derived a similar summation formula for the 2-dimensional case with a shifted center, we will follow their technique to derive the general d-dimensional summation formula.

2.1 A Summation Formula

(Adapted from [BCDL]): For convenience we will just consider the case where $\alpha = 0$. Let $\chi(x)$ be a characteristic function defined as follows

$$\chi(x, R) = \begin{cases} 1, & |x| \leq R \\ 0, & |x| > R \end{cases}$$

We now write:

$$N_0(R) = \sum_{n \in \mathbb{Z}^d} \chi(n, R).$$

By the Poisson summation formula,

$$\sum_{n \in \mathbb{Z}^d} \chi(n, R) = \sum_{n \in \mathbb{Z}^d} \tilde{\chi}(2\pi n, R),$$

where $\tilde{\chi}$ is the Fourier transform of χ. We calculate:

$$\tilde{\chi}(\xi, R) = \int_{\mathbb{R}^d} \tilde{\chi}(\xi, R) e^{-2\pi i \langle x, \xi \rangle} dx = \int_{|x|<R} e^{-2\pi i \langle x, \xi \rangle} dx,$$

where $\langle \cdot, \cdot \rangle$ indicates the inner product of x and ξ. By using the fact that χ is radially symmetric and a spherical coordinate substitution, we can rewrite this integral using the Bessel functions

$$\int_{|x|<R} e^{-2\pi i \langle x, \xi \rangle} dx = \frac{(2\pi)^{d/2} R^{d/2}}{|\xi|^{d/2}} J_{d/2}(|\xi| R),$$

Where J_v is the Bessel function of the first kind. Noting

$$\tilde{\chi}(0, R) = \int_{|x|<R} dx = \mathrm{Vol}(R, d),$$

We have

$$N_0(R, d) = \sum_{n \in \mathbb{Z}^d} \tilde{\chi}(2\pi n, R) = R^{d/2} \sum_{n \neq 0} \frac{J_{d/2}(2\pi |n| R)}{|n|^{d/2}} + \mathrm{Vol}(R, d).$$

Using a well-known asymptotic for the Bessel functions (See [GF])

$$J_v(t) = \sqrt{\frac{2}{\pi t}} \cos(t - \frac{1}{2} v\pi - \frac{1}{4}\pi) + O(t^{-3/2}),$$

We obtain:

$$\begin{aligned} N_0(R, d) &= R^{d/2} \sum_{n \neq 0} \frac{J_{d/2}(2\pi |n| R)}{|n|^{d/2}} + \mathrm{Vol}(R, d) \\ &= \pi^{-1} R^{\frac{d-1}{2}} \sum_{n \neq 0} |n|^{-\frac{d+1}{2}} \cos(2\pi |n| R - \frac{d+1}{4}\pi) + \mathrm{Vol}(R, d) + O\Big(\sum_{n \neq 0} |n|^{-\frac{d+3}{2}} R^{\frac{d-3}{2}} \Big). \end{aligned}$$

Now we can consider the normalized error term to get

$$\begin{aligned} F_0(R, d) &= R^{-\frac{d-1}{2}} \big(N_0(R, d) - \mathrm{Vol}(R, d) \big) \\ &\approx \pi^{-1} \sum_{n \neq 0} |n|^{-\frac{d+1}{2}} \cos(2\pi |n| R - \frac{d+1}{4}\pi). \end{aligned}$$

This is the summation formula that we can use to obtain results on the distribution of the error terms. To get a similar formula for the primitive case, we use the Möbius inversion formula.

2.2 Möbius Inversion formula

The classic version states that if g and f are arithmetic functions (i.e. functions from the set of positive integers to a subset of the complex numbers) satisfying

$$g(n) = \sum_{k|n} f(k) \quad \text{for every integer } n \geq 1$$

then

$$f(n) = \sum_{k|n} \mu(k)\, g\Big(\frac{n}{k}\Big) \quad \text{for every integer } n \geq 1.$$

Take $d = 2$ case as an example:

$$\begin{aligned}
\sum_{\substack{n_1^2+n_2^2 \leq R^2 \\ \gcd(n_1,n_2)=1}} 1 &= \sum_{n_1^2+n_2^2 \leq R^2} \sum_{k|\gcd(n_1,n_2)} \mu(k) \\
&= \sum_{k \leq R} \mu(k) \sum_{n_1^2+n_2^2 \leq R^2/k^2} 1 \\
&= \sum_{k \leq R} \mu(k) \sum_{n \leq R^2/k^2} r_2(n),
\end{aligned}$$

where $r_2(m)$ denotes the number of ways to express an integer m as the sum of two squares, i.e.

$$r_2(m) = \#\{(n_1, n_2) \in \mathbb{Z}^2 : n_1^2 + n_2^2 = m\}.$$

With the Hardy–Voronoi Identity [GH], we have

$$\sum_{n<R^2} r_2(n) \sim \pi R^2 - 1 + \frac{\sqrt{R}}{\pi} \sum_{m=1}^{\infty} \frac{r_2(m)}{m^{3/4}} \cos(2\pi\sqrt{m}R - \frac{3\pi}{4}).$$

Hence, the number of the primitive lattice points can be asymptotically expressed as

$$\sum_{k \leq R} \mu(k) \Big(\pi(\frac{R}{k})^2 - 1 + \frac{1}{\pi}\sqrt{\frac{R}{k}} \sum_{m=1}^{\infty} \frac{r_2(m)}{m^{3/4}} \cos(2\pi\sqrt{m}\frac{R}{k} - \frac{3\pi}{4}) \Big).$$

2.3 Statistical results

This next section documents the project's computational programs and the distribution results they produce. It covers two programs: a mature two-dimensional counter for the disk and a three-dimensional counter for the ball that is still in development. The section will describe the algorithms, their complexity, their efficiency, and what the diagnostic plots show.

Notation. For convenience we will take up different notation for this section, letting $N(r)$ be the number of points inside a ball of radius r, and $V(r)$ to denote the visible (primitive) lattice points. $E(r)$ and $E_{\text{prim}}(r)$ will denote the standard and primitive errors respectively. Unless otherwise stated, all counting functions and error terms will be centered at the origin. Additionally, when moving into higher dimensions, we will let a subscript denote the dimension, e.g., $E_3(r)$ will be the error for dimension $d = 3$.

2.4 The two-dimensional program

The counting core was rewritten once during development. This rewrite is the most consequential decision in the two-dimensional program.

Rewrite from enumeration to identities. Through version 6 the program counted by enumeration. It swept the disk's bounding box (or, across a radius sweep, only the annulus between successive radii), testing each integer point for membership and coprimality. This is simple and correct, but its time and memory both scale with the area swept, $O(r^2)$ per radius, which capped usable radii at a few times 10^4. Version 7 replaced enumeration with two classical identities that do not touch individual points. For the total count, a column-sum identity uses the disk's four-fold symmetry,

$$N(r) = 1 + 4r + 4\sum_{m=1}^{r} \lfloor \sqrt{r^2 - m^2} \rfloor,$$

evaluated in `count_M` as a function $M(R)$ of the integer squared radius $R = r^2$, so no square root of a radius is taken. Each floor is computed from a double-precision estimate and snapped to the exact integer with two comparisons, which keeps the count exact past $r^2 = 2^{53}$, where doubles can no longer represent every integer. This gives $N(r)$ in $O(r)$ operations instead of $O(r^2)$.

For the primitive count, `count_in_disk` uses Möbius inversion. Every lattice point is a unique positive-integer multiple of a primitive one, hence

$$V(r) = \sum_{d=1}^{r} \mu(d)\Big(M\big(\lfloor r^2/d^2 \rfloor\big) - 1\Big).$$

where $M(\lfloor r^2/d^2 \rfloor)$ counts the lattice points with both coordinates divisible by d, and the -1 removes the origin, which divides every d but is not itself primitive. A naive evaluation would call M at r arguments, but $\lfloor r^2/d^2 \rfloor$ takes only $O(r^{2/3})$ distinct values over $d = 1, \ldots, r$. The code walks d in blocks where that argument is constant, making one M-call per block weighted by the block's μ-sum. The Möbius function is sieved once at startup by `sieve_Möbius` (two passes, one signed byte per value).

Effect on resources. The effect is decisive. Per-radius time drops from $O(r^2)$ to $O(r)$ at the origin, and memory from $O(r^2)$ (the bounding box) to $O(r_{\max})$ (the Möbius table – tens of megabytes even at $r \sim 5 \times 10^7$). Accessible radii rose from a few times 10^4 to about 5×10^8.

Arbitrary centers. Version 8 added a cluster-sampling mode (denser radii in chosen windows). Version 9 generalized the center from the origin to an arbitrary point. The origin keeps the fast column-sum-plus-Möbius path, while any other center, whether integer, rational, or irrational, uses a shifted variant (`count_d_divisible`, `count_N_shifted`, `count_V_shifted`, dispatched by `count_at_center`). For each d the shifted sum counts the points divisible by d in the *original* disk rather than rescaling the disk by d. Counting in the original coordinates avoids a floating-point trap of the rescaled form, where a column tangent to the scaled boundary can gain or lose its boundary point to a rounding error in the square root. The shifted path costs $O(r \log r)$ per radius, the extra $\log r$ from a d-sum that no longer collapses by symmetry as it does at the origin.

Testing and output. The routines were checked against OEIS A000328 (lattice points in an origin-centered disk), against brute force at small radii and at several integer, rational, and irrational centers, and by confirming that the shifted code reproduces the origin path exactly up to $r = 10^5$. The integer-corrected square root was verified against exact integer roots near $r = 2 \times 10^8$. Each run produces a spreadsheet (one row per radius: counts, error terms, normalized ratios) and 21 diagnostic plots (raw error terms, log-log plots against a $\sqrt{r}$ reference, ratio diagnostics, histogram-and-Q-Q pairs under several normalizations, a refined residual (below), a correlation scatter, moving RMS curves, and Fourier spectra). The results below draw on the histogram and Q-Q pairs.

2.5 Two-dimensional results

Six meaningful runs were performed, with each sampling one center across a window of radii. For every radius the program forms the normalized errors $E(r)/\sqrt{r}$ and $E_0^*(r)/\sqrt{r}$, and for each, a histogram with a fitted normal and a kernel-density estimate plus a quantile–quantile (Q–Q) plot against the normal. Table 1 collects the summary statistics and the Shapiro–Wilk p-value tests normality.

The primitive error is Gaussian-shaped at every center. The clearest pattern is that $E_{\mathrm{prim}}(r)/\sqrt{r}$ is consistent with a Gaussian in all six runs. Shapiro–Wilk does not reject normality at any center (p from 0.33 to 0.92), skewness stays within ± 0.09, the fitted normal and the kernel estimate nearly coincide, and the Q–Q points lie on the diagonal (Figure 4). Holding across centers of very different arithmetic character makes this the suite's strongest result.

This, however, does need care, as failing to reject normality is not the same as establishing it. The windows are finite, and the values are not independent draws but successive values of one deterministic function, so the test's independence assumption does not strictly hold. The result is empirical evidence consistent with a Gaussian limiting law for the primitive error, not a proof of one. And the asymmetry matters too, as a clear rejection (small p) is a far firmer conclusion than a failure to reject (large p), which can also reflect limited power.

Table 1: Summary statistics for the normalized error terms across the six runs. Means and skews are dimensionless. "SW p" is the Shapiro–Wilk normality p-value.

			$E(r)/\sqrt{r}$			$E_{\text{prim}}(r)/\sqrt{r}$		
Center	Radius ($\times 10^6$)	n	mean	skew	SW p	mean	skew	SW p
Origin $(0,0)$, uniform	10–50	2001	-3.258	-0.332	5×10^{-8}	-0.528	-0.062	0.53
Origin $(0,0)$, cluster	1–10	4590	-3.250	-0.262	8×10^{-11}	-0.457	-0.004	0.92
$(10^5, -10^5)$, integer	0.5–1.5	501	-3.218	-0.265	1.3×10^{-2}	-0.417	-0.086	0.46
$(\sqrt{2}, \sqrt{2})$	0.5–1.5	1001	$+0.597$	$+0.030$	0.27	$+0.117$	-0.050	0.57
(φ, φ), golden	1–3	1001	$+0.724$	-0.101	3.2×10^{-2}	$+0.247$	$+0.087$	0.33
Liouville (≈ 0.110)	0.5–1.5	1001	-0.534	-0.189	4.2×10^{-4}	$+0.255$	-0.026	0.56

The total error isn't Gaussian, shown most clearly at the origin. The total-count error behaves differently. At the origin, with the two largest samples, $E(r)/\sqrt{r}$ departs robustly from normality ($p \sim 10^{-8}$ and 10^{-11}) with a consistent negative skew (-0.26 to -0.33), and its Q–Q plot bends below the diagonal in the left tail – a heavier-than-normal lower tail (Figure 3). Off the origin, the departure is weaker and center-dependent, as the well-approximable Liouville center also rejects ($p \sim 4 \times 10^{-4}$, negative skew) while the badly-approximable quadratic irrationals $\sqrt{2}$ and the golden ratio sit close to normal ($\sqrt{2}$ not rejected at $p = 0.27$, the golden ratio only marginally). The origin samples are an order of magnitude larger, so their rejection is the most reliable. The off-origin runs have less power.

Location of total error dependent on center. The mean of $E(r)/\sqrt{r}$ is strongly center-dependent and far from zero: about -3.25 at the origin, but $+0.60$ for $\sqrt{2}$, $+0.72$ for the golden ratio, and -0.53 for Liouville. With standard errors of a few hundredths ($\sigma/\sqrt{n}$, σ of order one, and n in the hundreds to thousands), every mean is many standard errors from zero. Hence, mean and skew vary systematically with the center, which is consistent with a Diophantine dependence on how well the center is rationally approximated. This is plausibly a genuine effect, though confirming that is for further analysis.

$\sqrt{r}$ is the right normalizing scale. The suite also tries a $\log(r)$ normalization, and the contrast supports $\sqrt{r}$. Dividing the primitive error by $\log(r)$ leaves a residual trend and a strongly non-normal distribution (origin: $p \sim 3 \times 10^{-11}$), whereas $\sqrt{r}$ removes the trend and gives the Gaussian-consistent distribution above ($p = 0.92$, same run). The error settling into a stable distribution under $\sqrt{r}$ but not $\log(r)$ serves as evidence that $\sqrt{r}$ is the right order for the fluctuations.

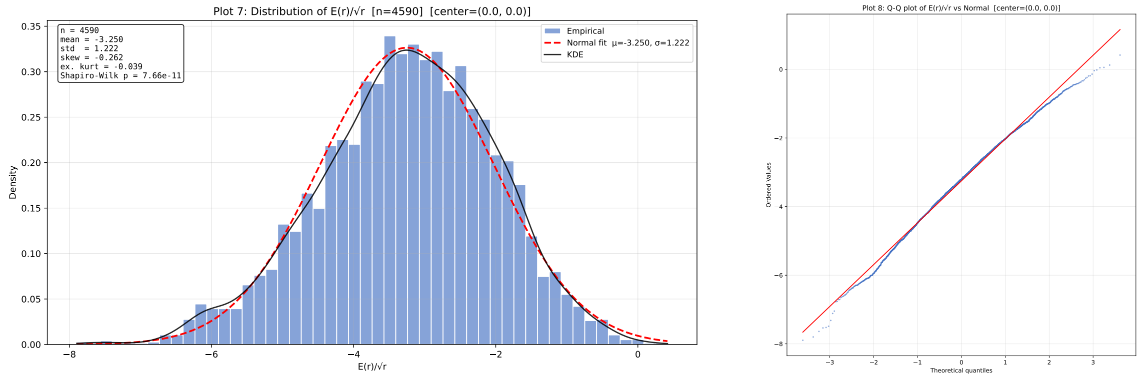
The refined residual reduces to the primitive error here. The suite includes a refined residual requested for the project,

$$F(r) = \Big(V(r) - \pi r^2 \sum_{d \le r} \mu(d)/d^2\Big)/\sqrt{r},$$

which subtracts the *truncated* Möbius form of the primitive main term rather than its asymptotic constant $6r^2/\pi$. Wherever computed, $F(r)$ is essentially identical to $E_{\text{prim}}(r)/\sqrt{r}$ (at the origin, mean and standard deviation -0.50, 3.58 against -0.53, 3.51). The refined object thus reduces to the plain primitive error at these scales. It is kept for completeness and for small-radius work, where the correction is visible.

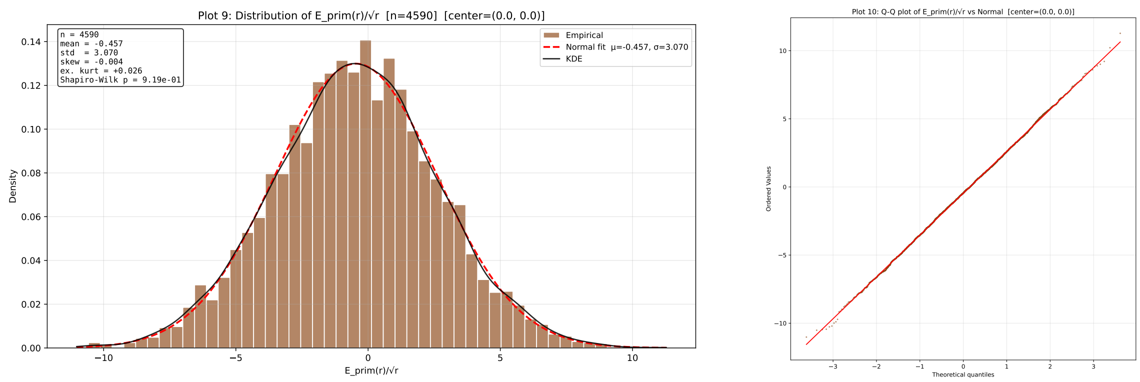
An integer-center consistency check. One run deserves a separate note. The center $(10^5, -10^5)$ is an integer point, and an integer translation maps the lattice onto itself, so $N(r)$ is identical to the origin's. The run does confirm it, as $E(r)/\sqrt{r}$ reproduces the origin distribution (mean -3.218 vs. -3.250, the same negative skew, the same rejection) over a different radius window. This illustrates the translation invariance of the total-count error and checks the shifted-center code, which correctly reduces to the original result. The primitive count is *not* translation-invariant (coprimality is defined on the global coordinates), so for E_{prim}, that center is a genuine extra data point.

The three-dimensional program The three-dimensional extension studies the same quantities for a ball, which is $N_3(r)$ lattice points and $V_3(r)$ primitive points inside a sphere of radius r. The main terms are the volume $\frac{4}{3}\pi r^3$ and $\frac{4}{3}\pi r^3/\zeta(3)$, where $1/\zeta(3) \approx 0.8319$ (the reciprocal of Apéry's constant) is the three-dimensional analogue of the planar density $1/\zeta(2) = 6/\pi^2$. According to the project's normalization, an error term in dimension d is divided by $r^{(d-1)/2}$. Hence, in three dimensions, r.



(a) (b)

Figure 3: Total-count error at the origin. The histogram (a) is left-skewed and the Q–Q plot (b) bends below the diagonal in the left tail. Shapiro–Wilk rejects normality ($p \sim 10^{-11}$, cluster run). The total error is not Gaussian.



(a) (b)

Figure 4: Primitive-count error at the origin. The histogram (a) matches its normal fit and the Q–Q plot (b) is straight. Shapiro–Wilk does not reject normality ($p = 0.92$, cluster run). The primitive error is Gaussian-consistent here, and at every center in Table 1.

Two facts shaped the implementation. First, the per-radius cost is far higher, as counting a ball is intrinsically three-dimensional, $O(r^3)$ per radius against $O(r)$ in the plane. Second, to keep the work fast, the counting moved to C++, with Python kept for orchestration and plotting. The C++ program reads a center and a list of radii and prints one row of counts per radius, which the existing Python plotting machinery consumes.

The primary direct method. The counting program uses direct enumeration. For a ball of radius r about (x_0, y_0, z_0), it walks the bounding box, tests each point against the ball with an exact squared-distance comparison, and tests coprimality of its global coordinates with a gcd. Planes and columns that are entirely outside the ball are skipped, and the innermost coordinate is bracketed by the chord half-width and confirmed point by point, so the count does not depend on square-root rounding. This is the primary path for two reasons: Firstly, it is transparently correct, and secondly it works for any center with no special handling, which is unlike the Möbius approach below that factors cleanly only at the origin. Its cost is $O(r^3)$ per radius. It has been spot-checked at small radii against brute force and against an independent Möbius implementation at the origin. A systematic test harness (e.g. against OEIS A005875, sums of three squares) and large-scale runs remain future work. The current version is single-threaded.

A faster origin-only alternative. A faster path exists for the origin. The plane's identities generalize: the total count slices into planar ones,

$$N_3(r) = M_3(r^2), \qquad M_3(R) = M(R) + 2\sum_{m=1}^{\lfloor\sqrt{R}\rfloor} M\big(R - m^2\big)\,,$$

reducing the three-dimensional count to the two-dimensional kernel M, and the primitive count follows by the same block-grouped Möbius inversion. This drops the per-radius cost to $O(r^2)$, which is a factor of r faster. Its drawbacks are why it is not primary, as the symmetry holds only at the origin, so any other center again needs a separate and more delicate treatment, and it reintroduces the Möbius sieve.

What the speedup is worth. The speedup matters less than it looks. Both methods are radius-limited: $O(r^2)$ and $O(r^3)$ both make a full sweep impractical at a finite radius, and the faster one mainly moves that ceiling from roughly 10^4 to 10^5 rather than opening a new range. Single-threaded, direct enumeration is comfortable to about $r = 10^3$ (seconds per radius) and reaches about 10^4 once parallelized. On memory it needs nothing beyond bookkeeping essentially since it streams through the points, whereas the Möbius path needs the $O(r_{\max})$ sieve. Since the three-dimensional work needs no large-radius results yet, the simpler and fully general direct method is the more appropriate primary choice, with the Möbius path held in reserve for large origin-centered runs.

3 Future Plans

3.1 Higher moments

Statistical evidence we gathered seems to suggest that the primitive case is Gaussian for $d = 2$. We hope to show this is the case analyzing the higher moments of the distribution term. This amounts to trying to understand the behavior of the following integral:

$$\int_0^T \left|F_0^*(R,2)\right|^k dR \qquad T \to \infty,$$

for some positive integer k. In order to show it is Gaussian we would show that for even k when weighted by the proper normalizing factor, the expression tends to a constant, and goes to 0 for odd k. Partial results in this direction come from [CCU] who showed:

$$\int \left|F_0^*(R,3)\right|^2 \gg R^2 \log R$$

and in the standard 3 dimensional case [HKLDW] showed:

$$\frac{1}{T^2 \log T}\int_0^T \left|F_0(R,3)\right|^2 dR \to C \qquad T \to \infty$$

For some constant C. We hope to extend the results in these papers.

3.2 Extending to higher dimensions

The three-dimensional pipeline is the main work remaining. A test harness mirroring the two-dimensional one (brute force at small radii plus a cross-check against the sum-of-three-squares sequence) would put the counts on the same footing as the plane. OpenMP over the outer coordinate, by which the loop is already structured for, should give about a six-fold speedup. Using the ball's symmetry about the origin gives another several-fold improvement, as the sign and permutation symmetries form a group of order 48 at the cost of careful bookkeeping on the symmetry boundaries. On the Python side, the orchestration and plotting suite need porting to three dimensions: the new constants $\frac{4}{3}\pi$ and $1/\zeta(3)$, the R rather than $\sqrt{R}$ normalization, and computing the truncated Möbius series for the refined residual in Python, as the direct counter has no sieve. The Möbius column-sum path can follow as an origin-only fast path if larger radii are wanted. The two-dimensional program is essentially complete and validated, and its natural extension is simply more radii and more complex sampling algorithms.

Code availability. The source for both programs is available at:
`https://github.com/RobinMa01/Gaussian-circle-sphere-normal-coprime-lattice-point-counter`

4 References

[AS] M. Abramowitz and I. A. Stegun (eds.), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, National Bureau of Standards Applied Mathematics Series 55, 1964.

[B99] P. M. Bleher, "Trace Formula for Quantum Integrable Systems, the Lattice Point Problem, and Small Divisors," in *Emerging Applications of Number Theory*, 1999.

[BCDL] P. M. Bleher, Z. Cheng, F. J. Dyson, and J. L. Lebowitz, "Distribution of the Error Term for the Number of Lattice Points Inside a Shifted Circle," *Communications in Mathematical Physics* **154** (1993), 433–469.

[CCU] F. Chamizo, D. Cristóbal, and A. Ubis, "Visible Lattice Points in the Sphere," *Journal of Number Theory* **131** (2011), 2137–2147.

[GF] G. B Folland, *Fourier Analysis and it's Applications*, 2th ed., 2009.

[GH] G. H. Hardy, *On the expression of a number as the sum of two squares*, Quarterly Journal of Pure and Applied Mathematics **46** (1915), 263–283.

[HL] G. H. Hardy and J. E. Littlewood, "Some Problems of Diophantine Approximation: Part II," *Acta Mathematica* **37** (1914), 193–239.

[Hux03] M. N. Huxley, "Exponential Sums and Lattice Points III," *Proceedings of the London Mathematical Society* **87** (2003), 591–609.

[HKLDW] T. Hulse, J. Kuan, D. Lowry-Duda, and A. Walker, "The Second Moment in the Generalized Gauss Circle Problem," *Forum Mathematicum* **27** (2015), 945–990.

[MV] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, Cambridge University Press, 2007.

[SS] E. M. Stein and R. Shakarchi, *Fourier Analysis: An Introduction*, Princeton University Press, 2003.

[Walfisz] A. Walfisz, *Gitterpunkte in Mehrdimensionalen Kugeln*, 1957.