1 Some History and General Remarks

The circle method was invented by Hardy and Littlewood during their common work with Sarinivasa Ramanujan in the years 1916-1917 on the asymptotic estimate of the Partition function \( p(n) \). The method gained its reputation when Hardy used it to give effective bounds for Waring's Problem (a.k.a Waring-Hilbert theorem). Later, works by Vinogradov, Kloosterman, Davenport and others modified the method to work in broader settings.

We'll present a modern approach of the method, which doesn't use complex analysis, based on [H-B]. See [MT] for a simpler treatment of Waring's problem with the circle method. See [IK] for a complete complex analysis based treatment.

2 How To Count By Integration

Let \( F(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n] \) be an irreducible form of degree \( d \) which defines a hypersurface by \( F = 0 \) in \( \mathbb{P}^{n-1} \).

Assume that \( R \) is a small box \( R = \prod_{i=1}^{n} (k_i, \lambda_i] \), define \( B_R = \{ x \in \mathbb{R}^n : x_i \in R \} \), and

\[
N_R(B) = |\{ x \in \mathbb{Z}^n \cap BR : F(x) = 0 \}| \quad (1)
\]

\( N_R(B) \) counts the number of rational points with bounded nominator and denominator on the hypersurface, note that one can extend that “point counting” function to any height function one wish to use.

We’ll be interested in asymptotic behavior of \( N_R(B) \) as \( B \to \infty \).

We’ll also be interested in p-adic distributions, for that, it’s sufficient to choose a vector \( (a_1, \ldots, a_n) \in \mathbb{Z}^n \) and \( m \in \mathbb{N} \) s.t. \( \gcd(m, a_1, \ldots, a_n) = 1 \) and modify the counting function

\[
N_R(B, m, a) = |\{ x \in \mathbb{Z}^n \cap BR : x = a (mod m) , F(x) = 0 \}|
\]

If \( m \) is composed of primes to high powers, then it’s describing the behavior in the adeles \( \mathbb{A}(\mathbb{Q}) \).

2.1 The Heuristic Principle

Let \( N(B) \) denote the number of integer vectors contained in \( BR \), easy counting argument shows that

\[
N(B) = (V(R) + o(1)) B^n \quad (2)
\]

Note that for \( x \in BR \), since \( F \) is of degree \( d \), \( F(x) = O(B^d) \). That means that the probability of \( F(x) = 0 \) for particular \( x \) is about \( B^{-d} \) in average.

So one should expect that \( N_R(B) \) should behave like \( N(B) B^{-d} = O(B^{n-d}) \).

2.2 The setup of the Hardy-Littlewood method

We’ll denote the function \( f(x) = e^{2\pi i \alpha x} \) by \( e_\alpha(x) \) for every \( \alpha \in \mathbb{R} \).

Recall the “Fundamental Theorem of Fourier Analysis”\(^4\)

\[
\int_0^1 e_\alpha(x) \, dx = \begin{cases} 
1 & x = 0 \\
0 & x \neq 0
\end{cases}
\]

\(^1\)Hardy established the asymptotics of \( p(n) \sim \exp(\pi \sqrt{\frac{2n}{3}}) \), and also asymptotic rate for the error. In 1937, Rademacher established the full formula.

\(^2\)Mainly, changing the focus from complex analysis to exponential sums via Fourier analysis.

\(^3\)Notice that argument makes sense only if \( n > d \).

\(^4\)In complex analysis, it corresponds to integrating over a circle, that’s how the circle method got its name originally.
Define the generating function $S(\alpha, BR) = S(\alpha) = \sum_{x \in \mathbb{Z}^n \cap BR} e_\alpha(F(x))$, for $\alpha \in [0, 1]$. Note that

$$N R(B) = \sum_{x \in \mathbb{Z}^n \cap BR} \int_0^1 e_\alpha(F(x)) = \int_0^1 S(\alpha) d\alpha$$

(3)

So we’ve effectively reduced the problem of rational point counting to evaluating (or at least obtaining asymptotics) for the integral on the RHS.

We’ll also define

$$S(\alpha, BR, m, a) = \sum_{x \in \mathbb{Z}^n \cap BR \atop x = a \mod m} e_\alpha(F(x)).$$

### 3 A Probabilistic View

Many ideas in modern number theory arose from probabilistic views\(^5\). Now we’ll try to give estimate the size of $S(\alpha)$ as $\alpha$ varies.

Assume that $-1 \leq k_i < \lambda_i \leq 1$, let $V$ denote the volume of $R$, $V(R) = \prod_i (\lambda_i - k_i)$.

Assume moreover that $0 < \alpha < 1$, one might expect that the numbers $e_\alpha(F(x))$ are randomly scattered on $[1, \alpha]$, as $x$ varies, for almost all $\alpha$, w.r.t Lebesgue measure, due to Weyl’s Equidistribution criterion.

If so, and assume that you take enough samples, the CLT assures you that $S(\alpha)$ should have some sort of “normal” behaviour.

Recall that the CLT also shows that the expected value should be about the size $\sqrt{N}$, when you take $N$ samples.

Therefore, according to the “square root cancelling philosophy”, we’ll get that $S(\alpha)$ is about the order of $B^{2\alpha}$, which is about the size of a random set chosen of $B^m$ elements. Combine it with the preceding argument, and you’ll expect that $n > 2d$ for the method to work.

That’s about the best there is, although you can combine the circle method with other facts for specialized cases and obtain results for fewer variables, i.e. Kloosterman achieved results for diagonal quadratics in four variables.

#### 3.1 When $e_\alpha$ changes its behavior?

Apparentely when $\alpha$ is near 0 (or equivalently 1), the function $S(\alpha)$ is indeed about the expected order of $B^{a-d}$, because $e_\alpha$ is almost constant, and one can verify that computationally. When $\alpha = 0$, the behavior is exactly like $B^n$ due to the counting argument mentioned before.

We might hope that when $\alpha$ is not too close to the endpoints of $[0, 1]$, $S(\alpha)$ will make a contribution of order $B^{2\alpha}$ on average.

one may see that such a simple distinction is not good enough.

Assume for instance that all the coefficients of $F$ are divisible by $q$, then we’ll actually see a behaviour similar to the behavior near 0 at fractions of the form $\alpha = \frac{a}{q}$, because we’ll get “artificial reduction” of $F$ there.

Define $S_q(\alpha) = \sum_{x \equiv (\mod q)} e_{\frac{a}{q}}(F(x))$, note that $e_{\frac{a}{q}}(F(x)) = e_{\frac{a}{q}}(F(y))$ if $y = x \mod q$,

and one gets

$$S\left(\frac{a}{q}\right) = \sum_{x \equiv (\mod q)} e_{\frac{a}{q}}(F(x)) \cdot \#\{y \in \mathbb{Z}^n \cap BR \atop y = x \mod q\}$$

And by simple counting argument

$$\#\{y \in \mathbb{Z}^n \cap BR \atop y = x \mod q\} = V(R) \frac{B^n}{q^n} + O(B^{n-1}),$$

so one get

$$S\left(\frac{a}{q}\right) = V(R) \frac{B^n}{q^n} S_q(\alpha) + O(\frac{B^n}{q^n})$$

Which means that at $\alpha = \frac{a}{q}$, if $S_q(\alpha) \neq 0$, the behavior is of order $B^n$, exactly like in $\alpha = 0$.

Due to the fact that $q$ is in the denominator, one would like to evaluate $S$ where $\alpha$ is close to fractions of small denominator.

---

\(^5\)i.e. The distribution of the primes and Cramer Model.
4 Arcs

Let \( \delta < \frac{1}{2} \), denote \( Q = B^\delta \), we’ll define \( I(a,q) = \left[ \frac{a}{q} - \frac{Q}{q^2}, \frac{a}{q} - \frac{Q}{q^2} \right] \) for every \( q \leq Q \), such that \( \gcd(a,q) = 1 \).

Note that those intervals are disjoint assuming \( B \gg 0 \).

We’ll define the Major Arcs \( \mathfrak{M} \) as

\[
\mathfrak{M} = \bigcup_{q \leq Q} \bigcup_{0 \leq a \leq q} I(a,q)
\]

(4)

We’ll define the Minor Arcs \( \mathfrak{m} = [0,1] \setminus \mathfrak{M} \).

Notice that length \( \length(I(a,q)) = \frac{2Q^\delta - d}{2Q^\delta} \), which means that the intervals \( I(a,q) \) are longer than \( B^{-d} \) by factor which tends to infinity.

Now we’ll approximate \( S(\alpha) \) on any interval \( I(a,q) \).

4.1 Major Arcs Approximation

Put \( \theta = \alpha - \frac{a}{q} \), and notice that if \( y = x \mod q \), then \( e_{\alpha}(F(y)) = e_{\frac{\theta}{q}}(F(x))e_{\theta}(F(y)) \).

So we get

\[
S(\alpha, BR, m, a) = \sum_{x \mod q} \sum_{y \in BR \atop x = a \mod m} e_{\alpha}(F(y))
= \sum_{x \mod mq} \sum_{y \in BR \atop x = a \mod m} e_{\frac{\theta}{q}}(F(x)) \sum_{y \mod mq} \sum_{y \in BR \atop y = x \mod mq} e_{\theta}(F(y))
\]

(5)

Note that the outer sum, is free of \( \alpha \). \( \alpha \) is found (implicitly) in \( \theta \) by \( \theta \)'s definition.

In the major arcs settings, \( |\theta| < 2B^{\delta-d} \), therefore we hope that the inner sum is well approximated by the integral

\[
I(B,\theta) = \int_{BR} e_{\theta}(F(r)) dr_1dr_2\cdots dr_n
\]

Actually, by calculation, one can deduce that

\[
\sum_{y \in BR \atop y = x \mod mq} e_{\theta}(F(y)) = (mq)^{-n}I(B,\theta) + O(B^{n-1-2\delta})
\]

(6)

We won’t bother to proof that (it’s very technical), but one should notice two things:

- The shape of the main term - \( (mq)^{-n}I(B,\theta) \).
- The exponential in the error term is strictly less than \( n \), because \( \delta \) is small.

Now plug in 6 in 5 and get

\[
S(\alpha) = \sum_{x \mod mq} \sum_{x = a \mod m} e_{\frac{\theta}{q}}(F(x)) \left( q^{-n}I(B,\theta) + O(B^{n-1+2\delta}) \right)
= \sum_{x \mod mq} \sum_{x = a \mod m} e_{\frac{\theta}{q}}(F(x)) \left( q^{-n}I(B,\theta) + O(q^nB^{n-1+2\delta}) \right)
\]

We’ll define \( S_q(a,m,a) = \sum_{x \mod mq} e_{\frac{\theta}{q}}(F(x)) \).

Integrate the equation from above over \( I(a,q) \) to get

\[
\int_{I(a,q)} S(\alpha) \, dx = S_q(a,m,a) (mq)^{-n}J(B) + O(Q^{n+1}B^{n-d-1+2\delta})
\]

\( ^6 \)Notice that \( \mathfrak{M} \subseteq [0,1] \), but because of periodicity of \( e_{\alpha} \), we can change the integral over any interval of length one, so we’ll disregard that for simplicity.
Where \( J (B) = \int_{Q^{n-d}} B^{n-d} I (B, \theta) \, d \theta. \)

Now sum over all the intervals \( I (a, q) \) and get the full integral over the Major Arcs,
\[
\int_{\mathbb{N}} S (\alpha) \, d \alpha = S (Q) J (B) + O \left( Q^{n+3} B^{n-d-1+2\delta} \right) = S (Q) J (B) + O \left( B^{n-d-1+(n+5)\delta} \right)
\]

Where \( S (Q) = \sum_{q \leq Q} \sum_{0 \leq a < q \atop (a, q) = 1} (mq)^{-n} S_q (a, m, a). \)

After some calculations, one can derive that \( J (B) \sim B^{n-d} c_\infty \) as \( B \to \infty \), where \( c_\infty \) is the singular integral
\[
c_\infty = \int_{w \in R, F(w) = 0} \frac{dw_2 \cdots dw_n}{\prod_{j=1}^{n} F (w)}
\]

Another way to define \( c_\infty \) is by \( c_\infty = \lim_{\epsilon \to 0} \left( 2\pi \right)^{-1} \int_{|F(x)| \leq \epsilon} 1 \, dx. \)

The singular integral can be thought as "real density" of the solutions \( F (x) = 0 \) in \( \mathbb{R} \) (notice that we haven’t required \( w \in BR \) there).

Now \( S (Q) \) captures the local density at any other valuation, apart from the infinity valuation.

We’ll define “\( p \)-adic density” \( c_p \) as \( c_p = \lim_{\epsilon \to 0} p^{-(a-1)k} \left| \{ x \pmod{p^k} \mid F (x) = 0 \pmod{p^k} \} \right| \),
then the singular series \( S (Q) \) becomes \( \prod_p c_p. \)

One can show that \( S_q (a, m, a) \) “factors” as a sort of product over different primes dividing \( q \), which can be thought like “Chinese Reminder Theorem”. If one defines
\[
S (q, m, a) = \sum_{0 \leq a < q \atop (a, q) = 1} S_q (a, m, a),
\]
then one actually get , due to \( S_q \) factorization ,
\[
S (q, m, a) = \prod_p S (p^j, p^{j'}, a).
\]

So we get \( S (Q) = \sum_{q \leq Q} S (q, m, a) = \sum_{q \leq Q} \prod_p S (p^j, p^{j'}, a) \), which can be thought as partial summation over corresponding Euler product. The main step from here is to take the infinite sum.

Estimating such sums in particular has been a major goal in number theory. I will remark that Deligne’s proof of the Riemann Hypothesis for Weil’s conjectures yields that if \( q \) is prime, and \( F \) is non-singular modulo \( q \) then \( S_q (a, m, a) \ll q^2 \), which is related to Hasse-Weil theorem and its generalizations over varieties - “Weil’s conjectures” which Tomer talked about a couple of weeks ago.

So assuming suitable estimates, which we won’t prove here,
one can actually get that \( S (Q) = m^{-n} \prod_p c_p + O (Q^{-n}). \)

Notice from the structure of the singular series, if the Hasse principle \textit{fails}, we won’t receive the expected value. However, if it \textit{works}, we will recieve the main term of \( (c_\infty \prod_p c_p) B^{n-d} \), which agrees with our heuristic principle, because, basically the product of the densities tells you what is the “probability” that a vector is indeed a local solution for all the local places determined by primes and the real numbers.

**4.2 Minor Arcs Estimate**

Our main goal here is to show that \( \int_{\mathbb{N}} S (\alpha, BR, m, a) \, d \alpha = O (B^{n-d}). \)

We’ll now impose some restrictions over \( F. \)

\textbf{First we’ll assume that \( F \) is a diagonal form, i.e. \( F (x) = c_1 x_1^d + \cdots + c_n x_n^d \).}

Notice that due to the exponent function properties, \( S (\alpha) \) factors as \( S (\alpha) = \prod_{\alpha^{\infty}} S_j (\alpha), \) where \( S_j (\alpha) \) are “one-dimensional counts”, \( S_j (\alpha, B, m, a_j) = \sum_{x \in \mathbb{Z} \cap [Bk_j, B\lambda]} e_\alpha \left( c_j x^d \right), \)
\[
x = a_j \pmod{m}
\]

One can then show, using Hölder inequality that \( \left| \int_{\mathbb{N}} S (\alpha) \, d \alpha \right| \leq \sup_{\alpha^{\infty}} \left| S_j (\alpha) \right|^{n-d} \int_{\mathbb{N}} \left| S_k (\alpha) \right|^2 \, d \alpha \), for some indices \( j, k. \) Note that \( \left| S_k (\alpha) \right|^4 = \sum_{x_1, x_2, x_3, x_4} e_\alpha \left( c_j \left( x_1^d + x_2^d - x_3^d - x_4^d \right) \right), \) so in the integration one gets the number of quadruples \( (x_1, x_2, x_3, x_4) \) in the specified region s.t. \( x_1^d + x_2^d = x_3^d + x_4^d \). One can actually relax the region constraints and count the number of quadruples in \( \mathbb{Z}^4 \cap [-B, B]^4. \)
We’ve actually reduced the problem to a whole different algebraic variety!
Now we’ll use the following number-theoretical lemma:

For every $\epsilon > 0$ there is a constant $c(\epsilon)$ s.t. $\# \{ m \in \mathbb{N} : m \mid N \} \leq c(\epsilon) N^\epsilon$ for every $N \in \mathbb{N}$

When $x_1^2 - x_2^2 > 0$, apply that to $N = x_1^2 - x_2^2$ and notice that $|x_1^2 - x_2^2|$ will be a positive divisor of $N$, and we get about $O (B^d)$ solutions. More over, if $x_1^2 = x_2^2$ then $x_1^2 = x_2^2$, and one gets $O (B^0)$ solutions. So in total, $\sum_i |S_i (\alpha)|^d \ll B^{2d} \epsilon$.

It remains to bound $S_j (\alpha)$ in order to bound the supremum in the RHS.

From now on, we’ll assume that $d = 2$.

Denote $\beta = \alpha c_j \in \mathbb{R} \setminus \mathbb{Q}$, $k = \kappa_j$, $\lambda = \lambda_j$, $a = a_j$, just to ease the notation.

Define $S_j (\alpha, B, m, a_j) = \sum_{B k \leq x \leq B \lambda} e_j \left( x^2 \right)$, which looks very similar to the statement of Weyl’s criterion. We’ll actually bound that in the similar way of deducing the criterion for polynomials of degree higher than 1. This method is known as “Weyl Differentiation”.

Look at $|S_j|^2$ for start, $|S_j|^2 = \sum_{|z| \leq \frac{1}{m} B} e_j (x^2 - y^2)$. Set $x = y + m z$, where $z$ goes over integer values in the range $|z| \leq \frac{\lambda (k - B)}{m}$. In particular we get that $|z| \leq 2 \lambda$.

Now substitute $y = a + m w$ and scale $w$ appropriately (i.e. $\mu \leq w \leq \nu$ for suitable $\mu, \nu$). The substitutions yield $x^2 - y^2 = m^2 z^2 + 2 am z + 2 m^2 w z$.

Then $|S_j|^2$ becomes $\sum_{|z| \leq 2 \lambda} e_j \left( m^2 z^2 + 2 am z \right) \sum_{\mu \leq w \leq \nu} e_j \left( 2 \beta m^2 w z \right)$, so one can get the trivial bound

$$|S_j|^2 \leq \sum_{|z| \leq 2 \lambda} \sum_{\mu \leq w \leq \nu} e_j \left( 2 \beta m^2 w z \right)$$ (8)

Now the inside sum is just summation over a geometric sequence, and one shows that $\sum_{\mu \leq w \leq \nu} e_j (2 \beta m^2 w z) \ll \min \left\{ B, \frac{1}{| \sin (\pi \beta m^2 z) |} \right\}$, so one get that $|S_j|^2 \ll \sum_{|z| \leq 2 \lambda} \min \left\{ B, \frac{1}{| \sin (\pi \beta m^2 z) |} \right\}$.

Note that near zero, sine is almost linear, so we’ll introduce the following notation $\varphi (t) = t - \max \{ z \in \mathbb{Z} : |z| \leq t + \frac{1}{2} \}$, due to the almost linearity, $\varphi (t) = \Theta (\sin (\pi t))$.

Define $Z_h = \{ z \in \mathbb{Z} : |z| \leq 2 B \frac{h}{\beta} \leq \varphi (2 \beta m^2 z) \leq \frac{h + 1}{\beta} \}$, where $|h| \leq 1 + \frac{B}{\beta}$.

If $|h| \geq 2$, then $\frac{1}{| \sin (\pi \beta m^2 z) |} \ll \frac{\beta}{|h|}$, and for $|h| \leq 1$, $B \ll \frac{B}{|h|}$, therefore, we can improve our bound $8$

$$|S_j|^2 \leq \left( \max_h \# Z_h \right) \sum_{|h| \leq 1 + \frac{1}{2}} \frac{B}{|h|}$$

Recall that $\sum_{h=1}^{B} \frac{1}{h} = \Theta (\log (n))$ and one can get the following bound

$$|S_j|^2 \leq \left( \max_h \# Z_h \right) \log B$$

Now one can show that $\# Z_h$ can be approximated by $\# Z_0$, and we get $d \ |S_j|^2 \ll \log B \# Z_0$.

Until now, we haven’t used the Minor Arcs at all.

Recall Dirichlet Diophantine Approximation theorem

Let $\alpha \in \mathbb{R}$, fix $N > 0$, then there exists a rational $\frac{a}{v}$ s.t. $|\alpha - \frac{a}{v}| < \frac{1}{N^2}, \ \gcd (u, v) = 1$, $v \leq N$

Using Dirichlet’s theorem one find a rational $\frac{a}{v}$ s.t. $|\alpha - \frac{a}{v}| < \frac{Q}{N^2}$ where $\gcd (u, v) = 1$, $v \leq Q B^{-1}$.

In particular, $|\alpha - \frac{a}{v}| \leq Q B^{-2}$, but because $\alpha \notin \mathbb{Q}$ then $v > Q$.

By direct calculation, one gets that $\varphi \left( \frac{2 m^2 c_j a z}{v} \right) \leq \frac{14 m^2 |a|}{B}$, so by definition of $\varphi$ one gets that $2 m^2 c_j a z = l \mod v$, $6$The sine appears because of the definition of the sine via Euler’s formula.

$7$Actually, $\gamma = \lim_{n \to \infty} \left( \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k} - \log (n) \right)$.

$8$From that one can also deduce a bound for the exponential sums over finite fields mentioned above, although it won’t be tight bound, it’ll be sufficient for our purpose here.
where $l$ satisfies $|l| \leq v^{1+4n^2/m^2} B^{-1}$.

So $l = O(1 + v B^{-1})$, for each $l$, the number of congruence classes determined by $l$ is finite $O(1)$, and the range $|z| \leq 2B$ contains at most $1 + 4Bv^{-1}$ numbers from each congruence class.$^9$

So after all that, $\# \{ z \in \mathbb{Z} \mid |z| \leq 2B, 0 \leq \phi(2\beta m^2 z) \leq B^{-1} \} \ll (1 + v B^{-1}) (1 + v^{-1} B) \ll 1 + v B^{-1} + B v^{-1}$. Now $v > Q = B^\delta$ therefore $v B^{-1} = B^{\delta - 1}, B v^{-1} = B^{1 - \delta}$, and

$\# \{ z \in \mathbb{Z} \mid |z| \leq 2B, 0 \leq \phi(2\beta m^2 z) \leq B^{-1} \} \ll B Q^{-1} = B^{1 - \delta}$.

So finally $|S_j|^2 \ll (B Q^{-1}) B \log B = B^{2 - \delta} \log B$, therefore $|S_j| \ll B^{1 - \frac{\delta}{2}} (\log B)^{\frac{1}{2}}$.

Combine it with the previous estimates $\int_m S(\alpha) d\alpha \ll \left( B^{1 - \frac{\delta}{2}} (\log B)^{\frac{1}{2}} \right)^{n-4} B^{2 + 2\epsilon}$. Assuming $\epsilon$ is small enough, we actually get $o(B^{n-2})$, when $n \geq 5$.

---

So we proved that for diagonal quadratic form of $n \geq 5$ variables, there is a constant $c(R, m, a)$ s.t. $N_R(B, m, a) = c B^{n-2} + o(B^{n-2})$ when $B \to \infty$.

References


$^9$To see that, substitute $z$ in the congruence equation for the available values, and due to the congruence, one has to divide by $v$. 

---

6