Counting Representations as a Sum of Two Squares: Circle Problem

We consider the number of representations \( r(n) \) of \( n \) as a sum of two integer squares,

\[
r(n) := \# \{(x, y) \in \mathbb{Z}^2 \mid n = x^2 + y^2 \}.
\]  

(1) We showed in class:

**Theorem 1.** The first moment

\[
S(X) := \sum_{n=1}^{X} r(n) = \pi X + O \left( \sqrt{X} \right)
\]  

(2) Here we consider the problem estimating the second moment of \( r(n) \).

**Theorem 2.** We have

\[
\sum_{1 \leq n \leq X} r(n)^2 = O(X \log X).
\]
**Proof.** We start from
\[ r(n)^2 = \# \{(x_1, x_2, y_1, y_2) \in \mathbb{Z}^4 \mid x_1^2 + x_2^2 = y_1^2 + y_2^2 = n \}. \]
Thus
\[ \sum_{1 \leq n \leq X} r(n)^2 \leq \# \{(x_1, x_2, y_1, y_2) \in \mathbb{Z}^4 \mid \max(|x_1|, |x_2|, |y_1|, |y_2|) \leq \sqrt{X}, x_1^2 + x_2^2 = y_1^2 + y_2^2 \}. \]
The right side counts the number of solutions of the Diophantine equation
\[(x_1 + y_1)(x_1 - y_1) = (y_2 + x_2)(y_2 - x_2).\]
with \(|x_i|, |y_i| \leq \sqrt{X}, (i = 1, 2)\). Now introduce new variables
\[ A := x_1 - y_1, \quad B := x_1 + y_1, \quad C := x_2 - y_2, \quad D := x_2 + y_2. \]
Then these variables are integers and solve the equation
\[ AB = CD \hspace{1cm} (5) \]
and we have
\[ |A|, |B|, |C|, |D| \leq 2\sqrt{X}. \hspace{1cm} (6) \]
Each solution \((x_1, x_2, y_1, y_2)\) to the original system corresponds 1-to-1 with distinct integer solutions \((A, B, C, D)\) of the system (5), (6), since we can invert the map
\[ 2x_1 = A + B, \quad 2y_1 = -A + B, \quad 2x_2 = C + D, \quad 2y_2 = -C + D. \]
Now set
\[ Y := 2\sqrt{X}, \]
and we have
\[ \sum_{n \leq X} r(n)^2 \ll \# \{(A, B, C, D) : 0 \leq A, B, C, D \leq Y, AB = CD \} \]
where there is a factor of 16 hidden in the symbol \(\ll\), which comes from imposing the restriction to nonnegative variables. (That is, there are 16 possible sign patterns for the numbers \((A, B, C, D)\) and it suffices to count the positive ones and multiply by 16.)

In making this change of variables we lost some information, since the new system may have additional integer solutions, corresponding to certain rational solutions of the first system. Thus we only hope to obtain an upper bound for \(\sum_{n \leq X} r(n)^2\) by
this approach (and the true number of solutions is somewhat smaller than our estimate.)

First we count the number of solutions \((A, B, C, D)\) to \(AB = CD\) having one variable equal to 0. Such solutions necessarily have two variables equal to zero, and there are \(Y = 2\sqrt{X}\) choices each for the other two variables, giving at most \(O(Y^2) = O(X)\) such solutions in all.

Next consider those solutions \((A, B, C, D)\) to \(AB = CD\) having all four variables nonzero. For such a solution we can factorize

\[
A = f_1f_2, \quad B = g_1g_2, \quad C = f_1g_1, \quad D = f_2g_2
\]

by taking in order

\[
f_1 := (A, C), \quad f_2 := \frac{A}{f_1}, \quad g_1 := \frac{C}{f_1}, \quad g_2 := \frac{D}{f_2},
\]

and this construction forces \((f_2, g_1) = 1\). From this fact, using \(Y = \sqrt{X}\) we obtain

\[
\sum_{1 \leq n \leq X} r(n)^2 = \sum_{1 \leq g_1 \leq Y} \left( \sum_{1 \leq f_1 \leq \min(f_1^{-1}, g_1^{-1})} 1 \right) + O(X)
\]

\[
\leq \sum_{1 \leq f_2 \leq g_1 \leq Y} \left( \sum_{1 \leq f_1 \leq f_2} 1 \right) + \sum_{1 \leq g_1 \leq f_2} \left( \sum_{1 \leq f_1 \leq \frac{Y}{f_2}} 1 \right) + O(X)
\]

\[
= \sum_{1 \leq f_2 \leq g_1 \leq Y} \left( \frac{Y}{g_1} + O(\frac{Y}{g_1}) \right) + \sum_{1 \leq g_1 \leq f_2} \left( \frac{Y}{f_2} + O(\frac{Y}{f_2}) \right) + O(X)
\]

\[
= 2Y^2 \log Y + O(Y) + O(X)
\]

\(\ll X \log X,\)

as asserted. \(\square\)

(3) Now define

\[
r_0(n) = \begin{cases} 
1 & \text{if } n = x^2 + y^2 \text{ for some integers } (x, y), \\
0 & \text{otherwise.}
\end{cases}
\]

(7)

Thus \(r_0(n)\) counts numbers represented as sums of two squares without multiplicity. We apply the two results above to show:
Theorem 3. There is a positive constant $C$ such that for $n \geq 1$,
\[ \sum_{n \geq x} r_0(n) \geq C_1 \frac{x}{\log x}. \] (8)

Remark. We may rewrite this assertion in Vinogradov notation as
\[ \sum_{n \geq x} r_0(n) \gg \frac{x}{\log x}. \]

Proof. We use the Cauchy-Schwartz inequality, which says that for real $A_j, B_j$
\[ \left( \sum_{j=1}^{n} (A_j B_n)^2 \right)^2 \leq \left( \sum_{j=1}^{n} A_j^2 \right) \left( \sum_{j=1}^{n} B_j^2 \right) \]
Here we take $A_n = r(n), B_n = r_0(n)$, and note that
\[ r(n)r_0(n) = r(n) \]
since $r(n) > 0$ if and only if $r_0(n) = 1$. Thus we obtain
\[ \left( \sum_{n \leq X} r(n) \right)^2 = \left( \sum_{n \leq X} r(n)r_0(n) \right)^2 \leq \left( \sum_{n \geq x} r(n)^2 \right) \left( \sum_{n \geq x} r_0(n)^2 \right). \]
We next note that since $r_0(n)$ takes the values 0 and 1
\[ r_0(n)^2 = r_0(n) \]
so that $\sum_{n \leq x} r_0(n)^2 = \sum_{n \leq x} r_0(n)$. Thus we obtain
\[ \sum_{n \leq X} r_0(n) = \sum_{n \leq x} r_0(n)^2 \geq \frac{\left( \sum_{n \leq X} r(n) \right)^2}{\sum_{n \leq X} r(n)^2} \]
To lower bound the numerator we can use Theorem 1, giving
\[ \sum_{n \leq X} r(n) \geq \pi X - O(\sqrt{X}) \geq 3X \]
for all sufficiently large $X$. By Theorem 2 we can upper bound the denominator
\[ \sum_{n \leq X} r(n)^2 \leq CX \log X. \]
Thus we obtain, for sufficiently large $X$
\[ \sum_{n \leq X} r_0(n) \geq \frac{(3X)^2}{CX \log X} \gg \frac{X}{\log X} \]
the desired result. $\square$.

Remark. The true order of magnitude of $\sum_{n \leq X} r_0(n)$ is known to be proportional to $\frac{X}{\sqrt{\log X}}$. 