Abstract. In this paper, we prove an asymptotic formula for the average number of solutions to the Diophantine equation $axy - x - y = n$ in which $a$ is fixed and $n$ varies.

1. Introduction

People have been considering Diophantine equations involving products and sums of some variables for a long time. The Diophantine equation

$$\prod_{i=1}^{k} x_i - \sum_{i=1}^{k} x_i = n$$

was studied by various people in the past few decades. It is easy to see that there always exists a few trivial solutions with most of $x_i$'s equal to 1. So people are asking about the number of solutions of this equation with all $x_i > 1$.

The case when $n = 0$ is very special, since it concerns the number of $k$-tuples with equal sum and product. In this case, it is conjectured by Misiurewicz [2] that $k = 2, 3, 4, 6, 14, 114, 174$ and $444$ are the only values of $k$ for which there are only trivial solutions. For general $n$, very little is known except that in 1970s Viola [6] proved that if $E_k(N)$ denotes the number of positive integers $n \leq N$ for which (1) is not soluble in integers $x_1, x_2, \ldots, x_k > 1$ then $E_k(N) = N \exp(-c_k(\log N)^{1-1/(k+1)})$ for some positive constant $c_k$. It is believed that for large $n$ equation (1) always has a nontrivial solution, which nevertheless is an open question in this area.

On the other hand, the case that $k = 3$ has received extensive attention, and several variations of this problem were studied. Brian Conrey asked whether the number of solutions in positive integers to the equation

$$xyz + x + y = n$$

2010 Mathematics Subject Classification. Primary 11D45, Secondary 11D09.

Key words and phrases. Diophantine equations, mean value theorem.
can be bounded by $O_\varepsilon(n^\varepsilon)$ for any $\varepsilon > 0$. Kevin Ford posed a generalisation of this problem, in which one would like to show that there are $O_\varepsilon(|AB|^{\varepsilon})$ nontrivial positive integer solutions to the equation

$$xyz = A(x + y) + B$$

for given nonzero $A, B \in \mathbb{Z}$.

In this paper, we consider another variation of the case that $k = 2$, namely the following equation

$$axy - x - y = n$$

where $a$ is a positive integer and $n$ is any nonnegative integer. This can be viewed as equation (3) in which $z$ is fixed and $A = 1$. Hence if the number of solutions of equation (4) is well understood, then one can probably understand the number of solutions of equation (3) simply by averaging over $a$.

Let

$$R_a(n) = \text{Card} \{(x, y) \in \mathbb{N}^2 : axy - x - y = n\}.$$

Here we are considering the number of positive integer solutions of equation (4) when $a$ is fixed and $n$ varies. A sharp asymptotic formula is established in this paper on the average of $R_a(n)$ over $n$. Notice the case that $a = 1$ is trivial, since then $R_1(n) = d(n+1)$ is just the divisor function of $n + 1$, the average of which is relatively well understood.

**Theorem 1.** For positive integers $a > 1$ and $N \geq 1$, we have

$$\sum_{0 \leq n \leq N} R_a(n) = \frac{1}{a} \left(N \log N - C(a)N\right) + \Delta_a(N)$$

where

$$C(a) = 2 \frac{\Gamma'(\frac{a-1}{a})}{\Gamma(\frac{a-1}{a})} + 2 \sum_{p|a} \frac{\log p}{p - 1} + \log a + 2\gamma + 1$$

and

$$\Delta_a(N) \ll \phi(a) \sqrt{\frac{N}{a}} \left(\log(aN)\right)^2.$$ (6)

Here $\Gamma(s) = \int_0^{+\infty} e^{-t} t^{s-1} dt$ is the standard $\Gamma$ function, and $\gamma$ is the Euler constant.

In fact, since the error term above is roughly of size $\sqrt{aN \left(\log(aN)\right)^2}$, it is conceivable that the main term will be inferior to the error term when $a \gg N^{1/2}$. So in order for the above asymptotic formula to really make sense, one would impose a condition on $a$, such as $a \ll N^{1/2}/\log N$.

Moreover, one can argue what is the right order of magnitude of the error $\Delta_a(N)$. In view of $R_1(n) = d(n+1)$, one can think $R_a(n)$ as a
“generalized” divisor function. Hence Theorem 1 just proves a mean value theorem for such a “generalized” divisor function. Since for the classical divisor function, the error is believed to be \( O(N^{1/4+\varepsilon}) \). It is very natural to pose such a conjecture for our error \( \Delta_a(N) \). The author suspects that following the van der Corput method on exponential sums as in the classical case, one can show \( \Delta_a(N) = O_a(N^{1/3-\delta}) \) for some \( \delta > 0 \).

Remark. It’s not hard to adapt the method in this paper in order to deal with equations like

\[ axy - bx - cy = n \]

and prove similar asymptotic formulas.

2. Preliminary Lemmas

We state several lemmas before embarking on the proof of Theorem 1. The content of Lemma 2 can be found, for example, in Corollary 1.17 and Theorem 6.7 of Montgomery & Vaughan [4], and Lemma 3 can be deduced from Theorem 4.15 of Titchmarsh [5] with \( x = y = (|t|/2\pi)^{1/2} \).

**Lemma 1.** When \( \sigma \geq 1 \) and \( |t| \geq 2 \), we have

\[
\frac{1}{\log |t|} \ll \zeta(\sigma + it) \ll \log |t|.
\]

**Lemma 2.** When \( 0 \leq \sigma \leq 1 \) and \( |t| \geq 2 \), we have

\[
\zeta(\sigma + it) \ll |t|^{1-\sigma} \log(|t|).
\]

**Lemma 3.** Let \( \chi \) be a non-principle character modulo \( a \) and \( s = \sigma + it \) and assume that \( t \in \mathbb{R} \). Then

\[
L(s, \chi) \ll \log(a(2 + |t|)), \text{ when } \sigma \geq 1
\]

and

\[
L(s, \chi) \ll (a|t|)^{\frac{1-\sigma}{2} + \varepsilon}, \text{ when } \frac{1}{2} \leq \sigma \leq 1.
\]

**Proof.** The first part follows from Lemma 10.15 of MV [4]. Now suppose that \( \chi \) is primitive. Then by Corollary 10.10 of MV [4],

\[
L(s, \chi) \ll (a|t|)^{\frac{1}{2} - \sigma} \log(a(2 + |t|))
\]

when \( \sigma \leq 0 \). Then by the convexity principle for Dirichlet series, for example as described in Titchmarsh [5] (cf. exercise 10.1.19 of MV [4]),

\[
L(s, \chi) \ll (a|t|)^{\frac{1}{2} - \sigma + \varepsilon}
\]
when $0 \leq \sigma \leq 1$. The proof is completed by observing that if $\frac{1}{2} \leq \sigma \leq 1$ and $\chi$ modulo $a$ is induced by the primitive character $\chi^*$ with conductor $q$, then

$$L(s, \chi) = L(s, \chi^*) \prod_{\substack{p \mid a \\ p \not| q}} (1 - \chi^*(p)p^{-s}) \ll |L(s, \chi^*)|2^{\omega(a)}.$$  

\[\square\]

**Lemma 4.** Let $T \geq 2$, then we have

$$\sum_{\chi \mod a} \int_{-T}^{T} |L(\frac{1}{2} + it, \chi)|^2 dt \ll \frac{\phi^2(a)}{a} T \log T.$$  

A proof of this lemma can be found for example in Montgomery [3].

**Lemma 5.** Let $a$ be a positive integer greater than 1 and $w > 0$, we have

$$\sum_{\substack{n \leq w \\ n \equiv -1 \mod a}} \frac{1}{n} = \frac{1}{a} \left( \log w - \frac{\Gamma'(\frac{a-1}{a})}{\Gamma(\frac{a-1}{a})} - \log a \right) + O(1/w).$$  

**Proof.** By Abel summation, the left hand side above is

$$\sum_{\substack{n \leq w \\ n \equiv -1 \mod a}} \frac{1}{n} = \left[ \frac{w + 1}{a} \right] \frac{1}{w} + \int_{1}^{w} \left[ \frac{t + 1}{a} \right] \frac{1}{t^2} dt$$  

$$= \frac{1}{a} + \int_{1}^{w} \frac{t + 1}{at^2} dt - \int_{1}^{w} \left\{ \frac{t + 1}{a} \right\} \frac{dt}{t^2} + O(1/w)$$  

$$= \frac{1}{a} \left( \log w + 2 - \int_{1}^{\infty} a \left\{ \frac{t + 1}{a} \right\} \frac{dt}{t^2} \right) + O(1/w).$$  

Recall that the digamma function $\psi(z)$ is defined as $\frac{\Gamma'(z)}{\Gamma(z)}$, and $\psi'(z)$ has a series expansion $\sum_{k=0}^{\infty} \frac{1}{(z+k)^2}$. So

$$\int_{1}^{\infty} \left( a \left\{ \frac{t + 1}{a} \right\} - \{t\} - 1 \right) \frac{dt}{t^2} = \sum_{h=0}^{\infty} \int_{0}^{a} \left( a \left\{ \frac{r + 1}{a} \right\} - \{r\} - 1 \right) \frac{dr}{(ah + r)^2}$$  

$$= \frac{1}{a^2} \int_{0}^{a} \left( a \left\{ \frac{r + 1}{a} \right\} - \{r\} - 1 \right) \psi'(\frac{r}{a}) \frac{dr}{r}$$  

(7)
Notice that

\[ a \left\{ \frac{r + 1}{a} \right\} - \{r\} - 1 = \begin{cases} 
0, & \text{if } 0 \leq r < 1 \\
1, & \text{if } 1 \leq r < 2 \\
\vdots & \vdots \\
a - 2, & \text{if } a - 2 \leq r < a - 1 \\
a - 1, & \text{if } a - 1 \leq r < a 
\end{cases} \]

Hence (7) is equal to

\[
\frac{1}{a^2} \left( \sum_{l=1}^{a-2} \left( \frac{l+1}{a} \right) \psi' \left( \frac{l+1}{a} \right) - \frac{l}{a} \psi \left( \frac{l}{a} \right) \right)
= \psi \left( \frac{a - 1}{a} \right) + \log a + \gamma
\]

The last equality follows from a well known property of the digamma function \( \psi \). Now the lemma is established after the observation \( \gamma = 2 - \int_1^\infty \frac{d}{t^2} \). □

Lemma 6. Let \( a \) be a positive integer greater than 1, then we have

\[
\frac{1}{\phi(a)} \sum_{\chi \neq \chi_0 \mod a} \bar{\chi}(-1) L(1, \chi) = \frac{-1}{a} \left( \frac{\Gamma' \left( \frac{a - 1}{a} \right)}{\Gamma \left( \frac{a - 1}{a} \right)} + \sum_{p \mid a} \frac{\log p}{p - 1} + \log a + \gamma \right).
\]

Proof. Let \( w \) be large compared to \( a \) (eventually we will let \( w \) goes to \( \infty \)). Then for non-principal characters \( \chi \mod a \), by Abel summation

\[
L(1, \chi) = \sum_{n \leq w} \frac{\chi(n)}{n} + O(a/w).
\]

Hence

\[
\frac{1}{\phi(a)} \sum_{\chi \neq \chi_0 \mod a} \bar{\chi}(-1) L(1, \chi)
= \frac{1}{\phi(a)} \sum_{\chi \neq \chi_0 \mod a} \bar{\chi}(-1) \sum_{n \leq w} \frac{\chi(n)}{n} + O(a/w).
\]
The main term on the right is
\[
\frac{1}{\phi(a)} \sum_{\chi \mod a} \bar{\chi}(-1) \sum_{n \leq w} \frac{\chi(n)}{n} - \frac{1}{\phi(a)} \sum_{n \leq w} \frac{1}{n}.
\]

We have
\[
\sum_{n \leq w} \frac{1}{n} = \sum_{m|n} \frac{\mu(m)}{m} \sum_{n \leq w/m} \frac{1}{n}
\]
\[
= \sum_{m|a} \frac{\mu(m)}{m} \left( \log(w/m) + \gamma + O(m/w) \right)
\]
\[
= \frac{\phi(a)}{a} \left( \log w + \sum_{p|a} \frac{\log p}{p-1} + \gamma \right) + O\left(\frac{d(a)/w}{a}\right).
\]

Here we are using the fact that
\[
- \sum_{m|a} \frac{\mu(m)}{m} \log m = \frac{\phi(a)}{a} \sum_{p|a} \frac{\log p}{p-1},
\]
this is because
\[
- \sum_{m|a} \frac{\mu(m)}{m} \log m = \sum_{p|a} \frac{\log p}{p} \sum_{k|a/p} \frac{\mu(k)}{k}
\]
\[
= \sum_{p|a} \frac{\log p}{p} \prod_{p'|a \atop p' \neq p} \left( 1 - \frac{1}{p'} \right)
\]
\[
= \sum_{p|a} \frac{\log p}{p} \left( \frac{1}{1 - \frac{1}{p}} \right) \prod_{p'|a} \left( 1 - \frac{1}{p'} \right)
\]
\[
= \frac{\phi(a)}{a} \sum_{p|a} \frac{\log p}{p-1}.
\]

On the other hand, we have
\[
\frac{1}{\phi(a)} \sum_{\chi \mod a} \bar{\chi}(-1) \sum_{n \leq w} \frac{\chi(n)}{n} = \sum_{n \leq w} \frac{1}{n}
\]
And by lemma 5, this is
\[
\frac{1}{a} \left( \log w - \frac{\Gamma'\left(\frac{a-1}{a}\right)}{\Gamma\left(\frac{a-1}{a}\right)} - \log a \right) + O(1/w).
\]
Thus we have shown that
\[
\frac{1}{\phi(a)} \sum_{\chi \neq \chi_0 \mod a} \bar{\chi}(-1)L(1, \chi)
\]
\[= \frac{1}{a} \left( \log w - \frac{\Gamma'(a-1)}{\Gamma(a-1/a)} \log a \right) - \frac{1}{a} \left( \log w + \sum_{p|a} \frac{\log p}{p-1} + \gamma \right) + O(a/w)
\]
\[= -\frac{1}{a} \left( \frac{\Gamma'(a-1)}{\Gamma(a-1/a)} + \sum_{p|a} \frac{\log p}{p-1} + \log a + \gamma \right) + O(a/w)
\]
Now the lemma is established when we let \(w \to \infty\) in the above. \(\square\)

3. Proof of Theorem

The starting point of the proof is the following observation. One can rewrite equation (4) in the following form
\[(ax - 1)(ay - 1) = an + 1.\] (8)
Namely we are going to count the following quantities,
\[R_a(n) = \text{Card}\{(x, y) \in \mathbb{N}^2 : (ax - 1)(ay - 1) = an + 1\}\]
and
\[S_a(N) = \sum_{0 \leq n \leq N} R_a(n).\]

After the change of variables \(u = ax - 1\) and \(v = ay - 1\), it follows that \(R_a(n)\) is the number of ordered pairs of natural numbers \(u, v\) such that \(uv = an + 1\) and \(u \equiv v \equiv -1 \mod a\).

Now the residue class \(u \equiv -1 \mod a\) and \(v \equiv -1 \mod a\) are readily isolated via the orthogonality of the Dirichlet characters \(\chi\ modulo a\). Thus we have
\[S_a(N) = \sum_{0 \leq n \leq N} \sum_{\substack{uv=an+1 \\mod a \\mod a \\mod a}} 1
\]
\[= \sum_{m \leq M} \sum_{\substack{uv=m \\mod a \\mod a \\mod a}} 1
\]
\[= \frac{1}{\phi^2(a)} \sum_{\chi_1 \mod a} \sum_{\chi_2 \mod a} \bar{\chi}_1(-1)\bar{\chi}_2(-1) \sum_{m \leq M} \sum_{uv=m} \chi_1(u)\chi_2(v),
\]
where \(M = aN + 1\).
Let
\[ a_m(\chi_1, \chi_2) = \sum_{uv=m} \chi_1(u)\chi_2(v). \]

Then we have
\[ S_a(N) = \frac{1}{\phi^2(a)} \sum_{\chi_1 \mod a} \sum_{\chi_2 \mod a} \overline{\chi}_1(-1)\overline{\chi}_2(-1) \sum_{m\leq M} a_m(\chi_1, \chi_2). \]

We analyze this expression through the properties of the Dirichlet series
\[ f_{\chi_1, \chi_2}(s) = \sum_{m=1}^{\infty} \frac{a_m(\chi_1, \chi_2)}{m^s} = L(s, \chi_1)L(s, \chi_2). \] (9)

This affords an analytic continuation of \( f_{\chi_1, \chi_2} \) to the whole complex plane.

By a quantitative version of Perron’s formula, as in Theorem 5.2 of MV [4] for example, we obtain
\[ \sum'_{m\leq M} a_m(\chi_1, \chi_2) = \frac{1}{2\pi i} \int_{\sigma_0-iT}^{\sigma_0+iT} f_{\chi_1, \chi_2}(s) \frac{M^s}{s} ds + R(\chi_1, \chi_2), \]
where \( \sigma_0 > 1 \) and
\[ R(\chi_1, \chi_2) \ll \sum_{4 \leq m < 2M} \min(1, \frac{M}{T|m-M|}) \sum_{m=1}^{\infty} \frac{|a_m(\chi_1, \chi_2)|}{m^{\sigma_0}}. \]

Here \( \sum' \) means that when \( M \) is an integer, the term \( a_M(\chi_1, \chi_2) \) is counted with weight \( \frac{1}{2} \).

Let \( \sigma_0 = 1 + \frac{1}{\log M} \). By (9) we have \( |a_m(\chi_1, \chi_2)| \leq d(n) \). Thus
\[ \sum_{m=1}^{\infty} \frac{|a_m(\chi_1, \chi_2)|}{m^{\sigma_0}} \ll \zeta(\sigma_0)^2 \ll (\log B)^2 \]
and so \( R(\chi_1, \chi_2) \ll \varepsilon M^{1+\varepsilon}T^{-1} \), for any \( \varepsilon > 0 \). Hence
\[ \sum_{m\leq M} a_m(\chi_1, \chi_2) = \frac{1}{2\pi i} \int_{\sigma_0-iT}^{\sigma_0+iT} f_{\chi_1, \chi_2}(s) \frac{M^s}{s} ds + O \left( \left( \frac{M}{T} + 1 \right) M^\varepsilon \right). \]

The error term here is
\[ \ll M^\varepsilon \]
provided that
\[ T \geq M. \]
The integrand is a meromorphic function in the complex plane and is analytic for all $s$ with $\Re s \geq \frac{1}{2}$ except for a possible pole of finite order at $s = 1$. Suppose that $T \geq 4$. By the residue theorem we have

$$
\frac{1}{2\pi i} \int_{\sigma_0 = \frac{1}{2} - iT}^{\sigma_0 + iT} f_{\chi_1, \chi_2} (s) \frac{M^s}{s} ds
$$

$$
= \frac{1}{2\pi i} \left( \int_{\sigma_0 = \frac{1}{2} - iT}^{\frac{1}{2} + iT} + \int_{\frac{1}{2} + iT}^{\sigma_0 + iT} \right) \frac{L(s, \chi_1) L(s, \chi_2) M^s}{s} ds
$$

$$
+ \text{Res}_{s=1} \left( L(s, \chi_1) L(s, \chi_2) \frac{M^s}{s} \right).
$$

Hence, by Lemmas 1, 2 and 3, the contribution from the horizontal paths is

$$
\ll (\log aT)^2 \frac{M}{T \log M} + \frac{(aT)^\varepsilon}{T} \int_{\frac{1}{2}}^{1} (aT)^{1-\sigma} M^\sigma d\sigma
$$

$$
\ll T^{-1} (aT)^\varepsilon M + T^{-1} (aT)^{1/2+\varepsilon} M^{1/2}
$$

and provided that $T \geq M^5$ this is

$$
\ll M^{-1}.
$$

On the other hand, the contribution from the vertical path on the right is bounded by

$$
M^{\frac{1}{2}} \sum_{2^k \leq T} 2^{-k} \int_{2^k}^{2^{k+1}} |L(\frac{1}{2} + it, \chi_1) L(\frac{1}{2} + it, \chi_2)| dt.
$$

And by Lemma 4

$$
\sum_{\chi_1 \chi_2 \mod a} \chi_1 (-1) \chi_2 (-1) \frac{1}{2\pi i} \int_{\frac{1}{2} - iT}^{\frac{1}{2} + iT} \frac{L(s, \chi_1) L(s, \chi_2) M^s}{s} ds
$$

$$
\ll M^{\frac{1}{2}} \sum_{2^k \leq T} 2^{-k} \int_{2^k}^{2^{k+1}} \left( \sum_{\chi \mod a} |L(\frac{1}{2} + it, \chi)| \right)^2 dt
$$

$$
\ll M^{\frac{1}{2}} \sum_{2^k \leq T} 2^{-k} \phi(a) \sum_{\chi \mod a} \int_{-2^{k+1}}^{2^{k+1}} |L(\frac{1}{2} + it, \chi)|^2 dt
$$

$$
\ll M^{\frac{1}{2}} \sum_{2^k \leq T} \phi^3(a) \frac{k}{a}
$$

$$
\ll \frac{\phi^3(a)}{a} M^{\frac{1}{2}} (\log M)^2
$$
on taking \( T = M^5 \).

Hence we obtain

\[
S_a(N) = \frac{1}{\phi^2(a)} \sum_{\chi_1 \mod a} \sum_{\chi_2 \mod a} \bar{\chi}_1(-1)\bar{\chi}_2(-1) \text{Res}_{s=1} \left( f_{\chi_1,\chi_2}(s) \frac{M^s}{s} \right) + \Delta_a(N)
\]

where

\[
\Delta_a(N) \ll \frac{\phi(a)}{a} \sqrt{M} (\log M)^2 \ll \phi(a) \sqrt{\frac{N}{a}} \left( \log(aN) \right)^2.
\] (10)

It remains to compute the residue at \( s = 1 \).

By (9) there are naturally two cases, namely

(i) \( \chi_1 = \chi_2 = \chi_0 \);

(ii) only one of \( \chi_1 \) and \( \chi_2 \) is equal to \( \chi_0 \) while the other one is equal to \( \chi \neq \chi_0 \).

In the latter case the integrand has a simple pole at \( s = 1 \) and the residue is

\[
\prod_{p | a} \left( 1 - \frac{1}{p} \right) L(1, \chi)(aN + 1) = \phi(a) L(1, \chi) N + \frac{\phi(a)}{a} L(1, \chi).
\]

By lemma 6, the sum over \( \chi \) for the second term above is small, hence can be absorbed in \( \Delta_a(N) \). While in the former case, the integrand has a double pole at \( s = 1 \) and the residue is

\[
\prod_{p | a} \left( 1 - \frac{1}{p} \right)^2 \left( M \log M - M \right).
\]

Hence we have shown that

\[
S_a(N) = \frac{1}{a^2} \left( (aN + 1) \log(aN + 1) - aN - 1 \right) + \left( \frac{2}{\phi(a)} \sum_{\chi \neq \chi_0 \mod a} \bar{\chi}(-1) L(1, \chi) \right) N + \Delta_a(N).
\]

Now by lemma 6, this is

\[
\frac{1}{a} \left( N \log N - C(a) N \right) + \Delta_a(N)
\]

where \( C(a) \) and \( \Delta_a(N) \) are given by (5) and (6) respectively.

This completes the proof of Theorem 1. \( \square \)
\[ axy - x - y = n \]

References

[1] R. K. Guy, *Unsolved problems in Number Theory*, second edition, Springer-Verlag, 1994.

[2] M. Misiołewicz, *Ungelöste Probleme*, Elem. Math., 21(1966) 90.

[3] H.L. Montgomery, *Topics in Multiplicative Number Theory*, Springer-Verlag, 1971.

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

[5] E.C. Titchmarsh, *The Riemann Zeta-Function*, 2nd edition, revised by D.R. Heath-Brown, Oxford, 1986.

[6] C. Viola, *On the diophantine equation* \( \prod_{i=0}^{k} x_i - \sum_{i=0}^{k} x_i = n \) *and* \( \sum_{i=0}^{k} \frac{1}{x_i} = \frac{a}{n} \), Acta Arithmetica, 1973, 22: 339-352.

Department of Mathematics, McAllister Building, Pennsylvania State University, University Park, PA 16802-6401, U.S.A.

E-mail address: huang@math.psu.edu